

Article

# 3-D Data Interpolation and Denoising by an Adaptive Weighting Rank-Reduction Method Using Multichannel Singular Spectrum Analysis Algorithm

Farzaneh Bayati <sup>1,\*</sup>  and Daniel Trad <sup>2</sup>

Department of Earth Science, University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada

\* Correspondence: farzaneh.bayati@ucalgary.ca

**Abstract:** Addressing insufficient and irregular sampling is a difficult challenge in seismic processing and imaging. Recently, rank reduction methods have become popular in seismic processing algorithms for simultaneous denoising and interpolating. These methods are based on rank reduction of the trajectory matrices using truncated singular value decomposition (TSVD). Estimation of the ranks of these trajectory matrices depends on the number of plane waves in the processing window; however, for the more complicated data, the rank reduction method may fail or give poor results. In this paper, we propose an adaptive weighted rank reduction (AWRR) method that selects the optimum rank in each window automatically. The method finds the maximum ratio of the energy between two singular values. The AWRR method selects a large rank for the highly curved complex events, which leads to remaining residual errors. To overcome the residual errors, a weighting operator on the selected singular values minimizes the effect of noise projection on the signal projection. We tested the efficiency of the proposed method by applying it to both synthetic and real seismic data.

**Keywords:** seismic data; interpolation; MSSA; SVD; rank reduction



**Citation:** Bayati, F.; Trad, D. 3-D Data Interpolation and Denoising by an Adaptive Weighting Rank-Reduction Method Using Multichannel Singular Spectrum Analysis Algorithm. *Sensors* **2023**, *23*, 577. <https://doi.org/10.3390/s23020577>

Academic Editor: Haitao Yu

Received: 23 October 2022

Revised: 15 December 2022

Accepted: 26 December 2022

Published: 4 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Seismic surveys are designed to keep a consistent grid of sources and receivers. In typical surveys, regularly sampled seismic surveys are uncommon or rare because of logistic obstacles or economic constraints. These limitations lead to large shot and receiver sampling intervals, which produce poorly, irregularly sampled seismic data along spatial coordinates with gaps without recorded traces. Seismic reconstruction methods can be divided into four main classes: signal processing-based methods, wave equation-based methods, machine learning-based methods, and rank reduction-based methods.

Most of the methods in the signal processing-based category are multidimensional and use prediction filters [1–3] and transform domains through means such as the Fourier transform [4–6], Radon transform [7,8], or Curvelet transform [9]. There are some hybrid techniques using combinations of the Fourier transform with prediction error filters [10] as a way to improve interpolation beyond aliasing. Vassallo et al. [11] proposed multichannel interpolation by matching pursuit (MIMAP), which interpolates high-order aliased data without making any prior information about the linearity of seismic events or the seismic wavefield model. Ghaderpour [12] proposed a multichannel anti-leakage least-squares spectral (MALLSSA) method that includes the spatial gradients of seismic data into the anti-leakage least-squares spectral (ALLSSA) to regularize seismic data beyond aliasing.

Wave-equation-based algorithms execute an implicit migration de-migration pair. Stolt [13] introduced a reconstruction algorithm that regularized a dataset by mapping it to physical space using migration–demigration. A finite-difference offset continuation filter was proposed by Fomel [14].

In recent years, machine learning has attracted much attention in geophysical studies. Deep learning (DL) is having a great influence on signal and image processing. However,

although very powerful, machine learning for geophysical problems is not mature yet, and applications often ignore the complications and subtleties of real data processing. Wang et al. [15] proposed a ResNet block design to interpolate anti-aliased seismic data. Siahkoochi et al. [16] proposed generative adversarial networks (GANs) to reconstruct heavily sub-sampled seismic data by giving up linearity and using an adaptive non-linear model. Convolutional neural networks (CNNs) are a novel strategy for the reconstruction of missing traces in pre-stack seismic images [17]. A new framework for training artificial neural networks (ANNs) was presented by Mikhailiuk and Faul [18] to restore corrupted multidimensional seismic signals. Wang et al. [19] utilized an eight-layer residual learning network (ResNet) with better back-propagation for interpolation.

The key aspect of rank reduction-based methods is that linear events in a clean seismic dataset are low-rank in the time and frequency domains. On the other hand, noise and missing traces increase the rank of data [20]. In the Fourier domain, rank reduction algorithms reduce the rank of Hankel/Toeplitz matrices generated from frequency slices. The singular spectrum analysis (SSA) method proposed by Sacchi [21], Tricket and Burroughs [22], Oropeza and Sacchi [23], and Bayati and Siahkoochi [24] works by rank-reduction of the Hankel matrix with an iterative algorithm in the frequency domain. Gao et al. [25] extended the SSA method to higher seismic data dimensions and called it a multichannel singular spectrum (MSSA). Interpolating regularly missing traces with a de-aliased MSSA method was proposed by Naghizadeh and Sacchi [26]. Kreimer et al. [27] developed the algorithm for pre-stack data via low-rank tensor completion. Kreimer and Sacchi [28] used the higher-order SVD for rank-reduction of the pre-stack seismic data tensor. Ely et al. [29] utilized a statistical test to control the complexity of regularization by tuning a single regularization parameter. Kumar et al. [30] developed rank-reduction techniques using matrix completion (MC) for seismic data reconstruction. Rekapali et al. [31] studied the application of multi-channel time slice singular spectrum analysis (MTSSSA) for 3D seismic data de-noising in the time domain and discussed the selection of the input processing window length. A damped OptShrink was studied by Siahshahr et al. [32], which selects the rank of the block of the Hankel matrix without knowing the final rank. Carozzi and Sacchi [33] proposed a method called (I-MSSA) that interpolates seismic data in its exact data coordinates that overcomes the problem of vertical errors in rough binning.

One of the necessary assumptions of rank-reduction-based techniques is that the Hankel matrix generated from clean and complete seismic data containing plane waves is low-rank, and its rank is equal to the number of plane waves present in the data. However, gaps and noise increase the rank, breaking this assumption. When dealing with data from complex geologies, rank selection becomes difficult and inconvenient. One approach to deal with this problem is to apply the algorithm to local windows. However, in most cases, it is not easy to find the proper window size because it is hard to decide if the structure in the local window is linear or not. Moreover, it is difficult to approximate the rank of each window. Choosing the wrong rank will lead to failure because, if overestimated, there will be a significant amount of residual noise remaining, and if underestimated, it will cause random noise and signal distortion.

In the presence of noise when the signal-to-noise ratio is low, data reconstructed by TSVD tend to contain a significant amount of residual noise. This residual noise is because part of the noise has a projection on the signal component projection. Nadakuditi [34] introduced an algorithm that provides a weighted approximation, where the weights reduce and threshold the singular values. Their method mitigates the effect of rank overestimation. If the rank is properly estimated, the algorithm will better estimate weak components of the signal subspace. Chen et al. [35] introduced a damping operator that shrinks the singular values containing significant particles of residual noise. Their method reached better results than the Cadzow rank-reduction method [22,23] in the reconstruction of highly noisy incomplete 5-D data. In this paper, we propose a method that automatically selects a rank for each local window. The method was proposed by Wu and Bai [36] for 2-D data. We improved it for 3-D data and propose an adaptive weighting rank reduction

method (AWRR) that selects the rank of the block of Hankel matrices automatically to reconstruct and denoise 3-D seismic data simultaneously. The strategy is to choose the second cutoff in the singular-value spectrum of the block Hankel matrix. We tested the efficiency of the proposed method by applying it to both synthetic and real seismic data.

## 2. Materials and Methods

Consider  $S(t, x, y)$  a block of 3-D seismic data in the  $t - x - y$  domain on  $N_t$  by  $N_x$  by  $N_y$  samples. ( $t = 1, \dots, N_t$ ), ( $x = 1, \dots, N_x$ ), ( $y = 1, \dots, N_y$ ). In the frequency domain, the data are represented as  $S(\omega, x, y)$  and ( $\omega = 1, \dots, N_\omega$ ). Each frequency slice of the data at a given frequency  $\omega_m$  can be represented by the following matrix:

$$S(\omega_m) = \begin{pmatrix} S(1,1) & S(1,2) & \dots & S(1,N_x) \\ S(2,1) & S(2,2) & \dots & S(2,N_x) \\ \vdots & \vdots & \ddots & \vdots \\ S(N_y,1) & S(N_y,2) & \dots & S(N_y,N_x) \end{pmatrix}. \quad (1)$$

To avoid notational confusion, let us ignore the argument  $\omega_m$ . Then, construct a Hankel matrix from each inline of  $S$ ; for the inline  $i$ th, the Hankel matrix in the  $m$ th frequency slice will be:

$$H_i = \begin{pmatrix} S(i,1) & S(i,2) & \dots & S(i,l) \\ S(i,2) & S(i,3) & \dots & S(i,l+1) \\ \vdots & \vdots & \ddots & \vdots \\ S(i,N_x-l+1) & S(i,N_x-l+2) & \dots & S(i,N_x) \end{pmatrix}. \quad (2)$$

At this step, we have  $N_y$  Hankel matrices of each inline. Then, the multichannel singular spectrum analysis (MSSA) constructs a block Hankel matrix  $M$  from Hankel matrices  $H_i$ :

$$M = \begin{pmatrix} H_1 & H_2 & \dots & H_n \\ H_2 & H_3 & \dots & H_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ H_{N_y-n+1} & H_{N_y-n+2} & \dots & H_{N_y} \end{pmatrix}. \quad (3)$$

The size of  $M$  is  $I \times J$ , where  $I = (N_x - m + 1)(N_y - n + 1)$  and  $J = mn$ . The integers  $m$  and  $n$  are chosen to make the Hankel matrices of  $H$  and the block Hankel matrix of  $M$  square matrices, or close to square. As we see, including the two spatial dimensions of the 3-D cube for each frequency can make the block Hankel matrix very large.

In seismic data, the observed data can be indicated as  $S_{obs} = R(S_0 + \eta)$ , where  $S_{obs}$  represents observed seismic data,  $S_0$  indicates the full noiseless data,  $\eta$  represents the random noise and the residuals, and  $R$  indicates the sampling matrix formed of zeros and ones. Using the MSSA algorithm, we can write the Hankel matrix of observed data as:

$$M = P + N, \quad (4)$$

where  $P$  represents low-rank Hankel matrix of the desired signal and  $N$  denotes the noise component which includes the gaps as well. In MSSA, the decomposition of the Hankel matrix in the rank-reduction step using TSVD will be as follows:

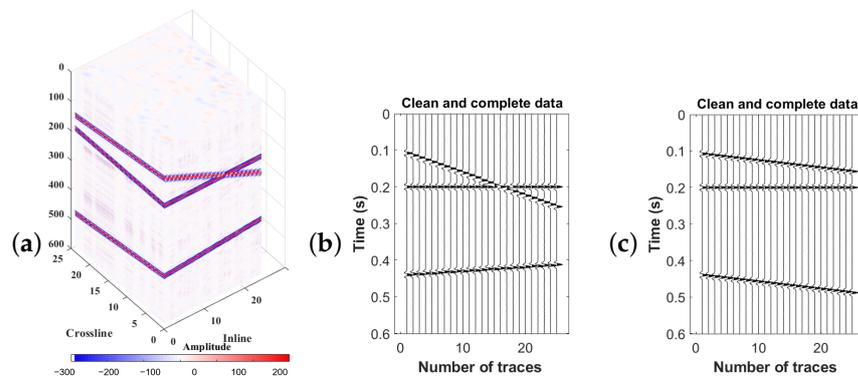
$$M = U\Sigma V^H. \quad (5)$$

If one knows the desired rank of the Hankel matrix, the estimated signal desired signal is recovered by:

$$\hat{M} = U_k \Sigma_k V_k^H, \quad (6)$$

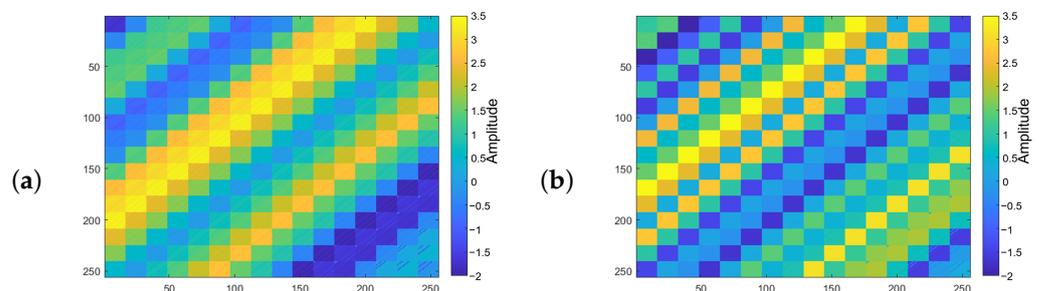
where  $\hat{M}$  is the estimated signal, and  $k$  is the predefined rank of the Hankel matrix equal to the number of the linear events in each local window.

Let us investigate the effect of linear events on the singular values in each frequency. Figure 1a shows the cube of 3-D data. Figure 1b indicates an inline of the data and Figure 1c represents a slice of data in a cross-line direction.



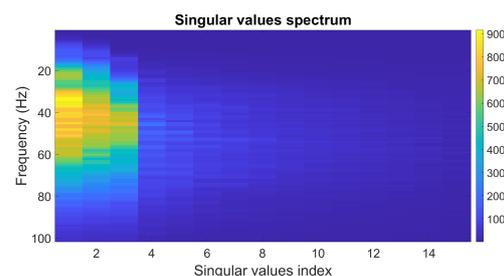
**Figure 1.** (a) Cube of 3-D data having three linear events. Slice of data in (b) inline direction, (c) cross-line direction.

We investigate the singular value distribution of the data by converting it to the  $(f, x, y)$  domain. Next comes generating the block Hankel matrix in each frequency slice. Figure 2a shows the block Hankel matrix for the frequency slice of 20 Hz. Figure 2b displays the block Hankel matrix generated in the frequency slice of 50 Hz. According to Figure 2a,b, it is clear that the block Hankel matrix for the frequency of 20 Hz is smoother than the one for the frequency of 50 Hz. It signifies that the block Hankel matrices for the higher frequencies require a higher rank than those in the lower frequencies to recover.



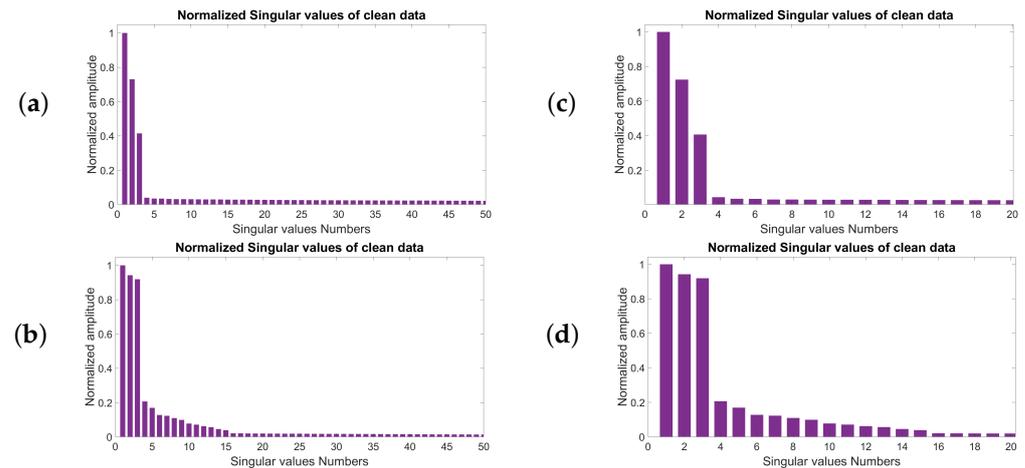
**Figure 2.** Block Hankel matrix for frequency slice of (a) 20 Hz, (b) 50 Hz. The color bar represents the absolute value of the FFT of input data in the mentioned frequency.

The singular-value spectrum of the data's block Hankel matrices for its frequency range is presented in Figure 3. Figure 3 represents the first 15 singular values' spectrum. This figure illustrates how the number of nonzero singular values of the block Hankel matrix at low frequencies is equal to the number of linear events but increases with frequency.



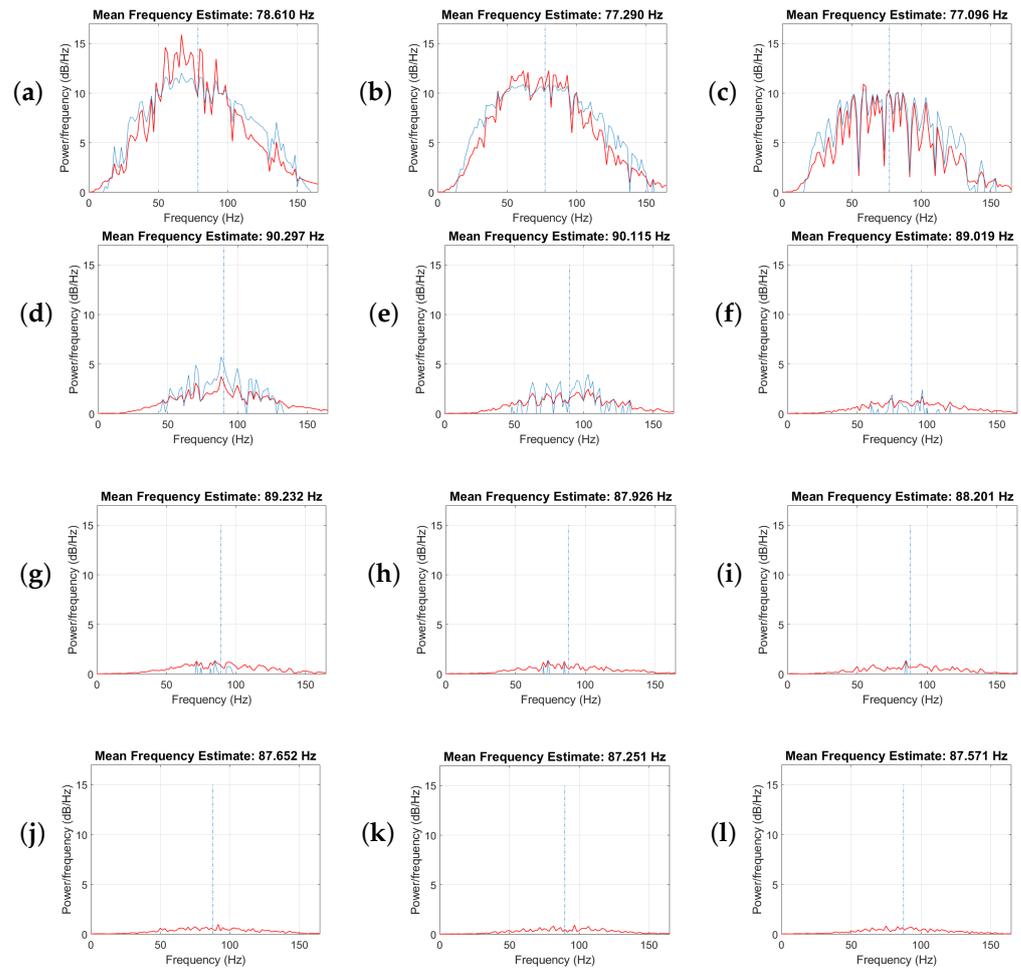
**Figure 3.** Singular-value spectrum. The color bar represents the magnitude of the singular values.

Similarly, Figure 4a represents the bar plot of the normalized singular values for the frequency slice of 20 Hz. Figure 4b shows the zoomed image for the first 20 singular values. Figure 4c shows the bar plot of the singular values for the frequency slice of 50 Hz. Figure 4d is the zoomed image for the first 20 singular values. From the figures, one can see that for the frequency of 20 Hz, the maximum difference in the energy between two adjacent singular values occurs in the third singular value, and it is the same as the number of linear events. However, for the frequency slice of 50 Hz, there is a second group of nonzero singular values. It means that we need to keep more singular values to recover all the frequencies completely.



**Figure 4.** Normalized singular value distribution of the block Hankel matrix of 3-D data with three linear events in the frequency slice: (a) 20 Hz, (b) 50 Hz. The first 20 singular values of the same data for frequency slice of (c) 20 Hz and (d) 50 Hz.

To analyze the effect of each singular value on the energy of each frequency for the data presented in Figure 1, data were decomposed into their singular matrices and recovered with each singular value from 1 to 15 separately. Then, the energy per frequency or power spectrum of the recovered data with each singular value is calculated. In Figure 5, the red line represents the graph of frequency energy per frequency, and the blue line shows the estimate of the mean normalized frequency of the power spectrum. Figure 5a is the result of recovering data with the first singular value. Figure 5b shows the energy per frequency for the recovered data with the second singular value. Figure 5c is the energy per frequency for the recovered data, including just the third singular value. The estimated mean frequency for these first three singular values is 77 Hz. Figure 5d–f represent the energy per frequency for the recovered data, including just the 4th, 5th, and 6th singular values, respectively. The estimated mean frequency for these singular values is 90 Hz. Figure 5g–i indicates the energy per frequency for the recovered data, including just the 7th, 8th and 9th singular values, respectively. The estimated mean frequency for these singular values is 88 Hz. Figure 5j–l show the energy per frequency for the recovered data, including just the 10th, 11th, and 12th singular values, respectively. It is clear that most energy of the useful signal is recovered with the first three singular values. Nevertheless, the shift in the mean frequency from 77 to 90 Hz states that there is leakage for the higher frequencies in the data. We conclude from this test that despite the remaining residual errors in the presence of the additive noise, we need to choose the rank of the block Hankel matrix to be greater than the number of linear events in each processing window of frequency and space.



**Figure 5.** Analysis of the power spectrum for each singular value. Panels (a–l) show the power spectrum recovered by the 1st to 12th singular values, respectively.

To find the best rank of data that minimizes residual errors, we tested different kinds of data containing various numbers of events with different slopes and amplitudes. The best result is when we choose the rank from the second cutoff of singular values instead of the first cutoff.

In all tested data, it is possible to see the first cutoff with a distinct change in the energy of singular values. However, sometimes there is no abrupt change in the amplitude of the singular values for the second cutoff, especially in the presence of a high level of random noise. Finding the second cutoff could be a challenge when the quality of the signal is poor. With a careful look at where the first and the second cutoff in clean and complete data happen, we conclude that there is a linear relationship between the first and second cutoff. The best estimate for the second cutoff can be estimated as follows:

$$\mathcal{T}(\Sigma, k) = \max_i \frac{\sigma_i^2}{\sigma_{i+1}^2}, \quad (7)$$

$$\mathcal{B}(\Sigma, \tilde{k}) = 3 \times \mathcal{T}(\Sigma, k), \quad (8)$$

where  $\sigma_i$  is the  $i$ th singular value of the Hankel matrix in each frequency.  $\mathcal{T}(\Sigma, k)$  indicates the operator that finds the rank at the point where the two following singular values become more scattered.  $\mathcal{B}(\Sigma, \tilde{k})$  indicates the second cutoff of the singular-value spectrum of the block of a Hankel matrix in each frequency slice.  $\tilde{k}$  can be introduced as the optimal rank of the block Hankel matrix that minimizes the Frobenius-norm difference between the approximated and the exact signal components.

Substituting Equation (8) to the rank-reduction step will give us:

$$\tilde{M} = \mathbf{U}_{\tilde{k}} \mathcal{B}(\boldsymbol{\Sigma}, \tilde{k}) \mathbf{V}_{\tilde{k}}^H, \tag{9}$$

By applying Equation (9) on the rank-reduction step of MSSA, the rank-reduced block Hankel matrix is reconstructed. The next step is averaging anti-diagonals of the recovered block Hankel matrix. It recovers the signal in the Fourier domain for each frequency slice. The rank-reduction step leads to gradually reconstructing the missing traces and can be used together with an iterative algorithm to replace the missing traces with the reconstructed traces. An iterative algorithm is a practical approach for random noise attenuation and seismic data amplitude reconstruction. The algorithm is written as follows:

$$\tilde{S}_{n+1}(\omega_m) = \alpha_n \mathbf{S}_{obs} + (\mathbf{I} - \alpha_n \mathbf{R}) \mathcal{F}^{-1} \mathcal{A} \mathcal{R} \mathcal{B} \mathcal{F} \left( \tilde{S}_n(\omega_m) \right), \quad n = 1, 2, \dots, N \tag{10}$$

where  $\omega_m$  is the temporal frequency,  $\mathbf{I}$  is a matrix of ones,  $\alpha_n \in (0, 1]$  is a weight factor, which decreases linearly with iterations,  $\mathcal{F}$  indicates the Fourier transform, and  $\mathcal{F}^{-1}$  shows the inverse Fourier transform.  $\mathcal{B}$  points to the Hankelization operator to generate the block Hankel matrix,  $\mathcal{R}$  reveals the rank reduction operator, and  $\mathcal{A}$  gives the averaging anti-diagonal operator.

Selecting the rank using Equation (9) recovers all the signal components. One of the advantages of this rank-reduction method is that it is adaptive and data-driven, and there is no need to set any parameter for the rank-reduction step in each processing window. Moreover, choosing the second cutoff instead of the first one leads to a better reconstruction of the high frequencies. However, selecting large ranks leads to more residual errors in the recovered data. That is why we need to apply a weighting operator to reduce the effect of noise on the projected signal components to recover higher frequencies.

We are looking for a weighting operator  $\hat{W}$  to adjust the singular values of  $\tilde{M}$  to calculate the best estimation of the desired signal. Nadakuditi [34] proposed an algorithm for low-rank matrix denoising that can be summarized as follows

1.  $M_{n \times m}$  is the signal plus noise Hankel matrix.
2.  $k$  is the best effective rank that can represent the signal.
3. For  $i = 1 : k$ ,

$$\text{compute } \hat{w}_i = \left( -\frac{2}{\sigma_i} \frac{\mathcal{D}(\sigma_i; \boldsymbol{\Sigma})}{\mathcal{D}'(\sigma_i; \boldsymbol{\Sigma})} \right), \tag{11}$$

4. End for loop,
5. Compute results as  $\hat{M} = \sum_{i=1}^k \hat{w}_i \sigma_i u_i v_i^H$

In this algorithm,  $\mathcal{D}(\sigma; \boldsymbol{\Sigma})$  is computed as:

$$\mathcal{D}(\sigma; \boldsymbol{\Sigma}) = \frac{1}{k} \text{Tr}(\sigma(\sigma^2 \mathbf{I} - \boldsymbol{\Sigma} \boldsymbol{\Sigma}^H)^{-1}) \frac{1}{k} \text{Tr}(\sigma(\sigma^2 \mathbf{I} - \boldsymbol{\Sigma}^H \boldsymbol{\Sigma})^{-1}) = \left[ \frac{1}{k} \text{Tr}(\sigma(\sigma^2 \mathbf{I} - \boldsymbol{\Sigma}^2)^{-1}) \right]^2, \tag{12}$$

where  $\mathcal{D}$  represents  $D$ -transform and  $\text{Tr}(\cdot)$  denotes the trace operator of the input.

The  $\mathcal{D}'$  represents the derivative of  $\mathcal{D}$  with respect to  $\sigma$ :

$$\begin{aligned} \mathcal{D}'(\sigma; \boldsymbol{\Sigma}) &= 2 \left[ \frac{1}{k} \text{Tr}(\sigma(\sigma^2 \mathbf{I} - \boldsymbol{\Sigma}^2)^{-1}) \right] \\ &= \left[ \frac{1}{k} \text{Tr}((\sigma^2 \mathbf{I} - \boldsymbol{\Sigma}^2)^{-1} - 2\sigma(\sigma^2 \mathbf{I} - \boldsymbol{\Sigma}^2)^{-2} \sigma) \right] \\ &= \frac{2}{k^2} \left[ \text{Tr}(\sigma(\sigma^2 \mathbf{I} - \boldsymbol{\Sigma}^2)^{-1}) \right] \left[ \text{Tr}((\sigma^2 \mathbf{I} - \boldsymbol{\Sigma}^2)^{-1} - 2\sigma^2(\sigma^2 \mathbf{I} - \boldsymbol{\Sigma}^2)^{-2}) \right], \end{aligned} \tag{13}$$

The  $D$ -transform describes how the distribution of the singular values of the sum of the independent matrices is related to the distribution of the singular values of the individual matrices [37]. Benaych and Nadakuditi [38] indicate that the principal singular values and vectors of a large matrix can be set apart as the singular values of the signal matrix and the  $D$ -transform of the limiting noise-only singular value distribution. From Equation (11), the weighting operator  $\hat{W}$  can be written as:

$$\hat{W} = \text{diag}(\hat{w}_1, \hat{w}_2, \dots, \hat{w}_k). \quad (14)$$

We can substitute the weighting algorithm obtained from Equation (14) into Equation (9) to enhance the results of the rank-reduction step as:

$$\hat{M} = \mathbf{U}_{\tilde{k}} \hat{W} \mathcal{B}(\Sigma, \tilde{k}) \mathbf{V}_{\tilde{k}}^H, \quad (15)$$

where  $\hat{M}$  indicates the reduced rank block Hankel matrix. Missing traces can be interpolated completely by applying the iterative algorithm. This adaptive-weighting rank-reduction (AWRR) method leads the way that sorts out the rank of the block Hankel matrix automatically while reducing the effect of the residual errors.

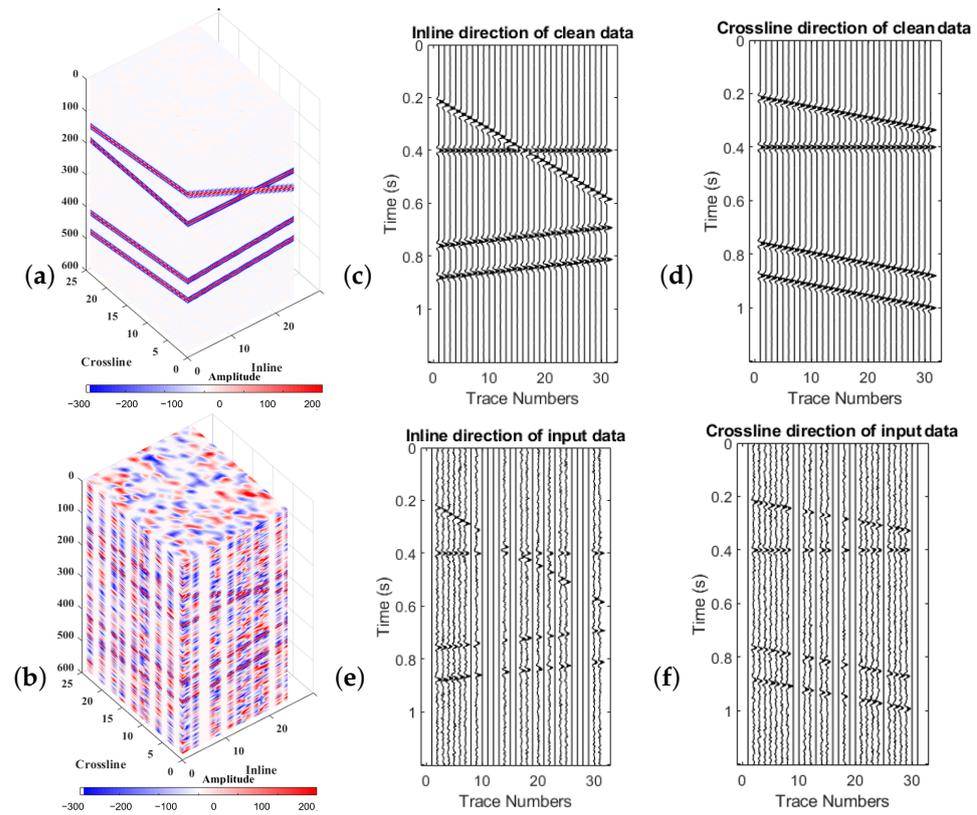
The weighting operator satisfies the equation below [34]:

$$\|\hat{M} - P\| \leq \epsilon. \quad (16)$$

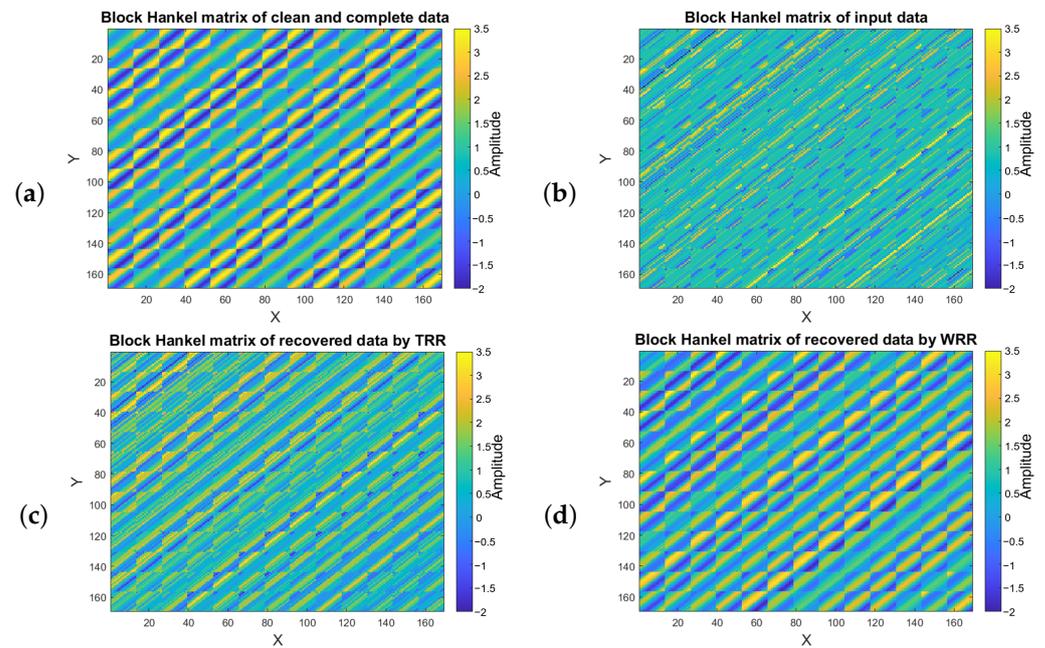
To understand the effect of the weighting operator on the singular values of the block Hankel matrices, the predefined rank-reduction method (TRR) and the weighting rank-reduction method (WRR) with predefined rank were applied to a block Hankel matrix. Figure 6a shows clean and complete 3-D synthetic data having four linear events. Figure 6b indicates the same 3-D data with  $SNR = 2$  and 51% missing traces. Figure 6c,d show slices of data in inline directions for the cubes of Figure 6a and Figure 6b, respectively. Figure 6e,f are the slices of the cubes of Figure 6a and Figure 6b in the crossline direction, respectively.

Figure 7 shows the Hankel matrices of data in Figure 6 for the frequency slice of 20 Hz. Figure 7a is the block Hankel matrix of clean and complete data. Figure 7b corresponds to the block Hankel matrix of data with  $SNR = 2$  and 51% missing traces. Figure 7c shows the block Hankel matrix after applying the predefined rank = 12 for ten iterations. Figure 7d indicates the block Hankel matrix after applying the weighting operator to adjust the singular values of the data using TSVD with predefined rank = 12 after ten iterations. We can see from Figure 7 that the block Hankel matrix of the WRR method is smoother than the result of applying a predefined rank = 12 after ten iterations.

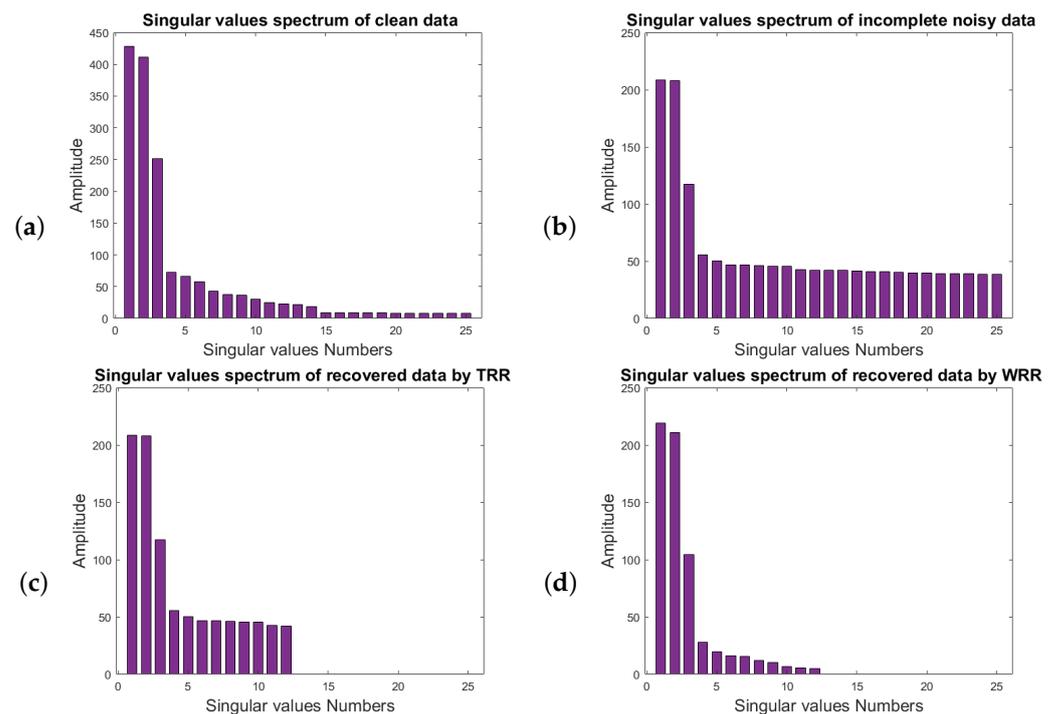
Figure 8 corresponds to the spectrum of the first 25 singular values of the block Hankel matrices of Figure 6. Figure 8a represents the singular-value spectrum of the block Hankel matrix of clean and complete data for the frequency of 20 Hz, the first three singular values indicate the useful signal that relates to the coherent events in the data. Figure 8b relates to the singular-value spectrum of the noisy and incomplete data, Figure 8c shows the singular-value spectrum of the noisy and incomplete data after applying TSVD with rank = 12. Figure 8d indicates the singular-value spectrum of the noisy and incomplete data after applying TSVD with rank = 12 and the weighting operator. We can see that the singular spectrum of the data after applying the weighting operator is much closer to the spectrum of the desired signal in Figure 8a.



**Figure 6.** (a) Cube of 3-D synthetic clean and complete data; (b) Cube of 3-D synthetic data with  $SNR = 2$  and 51% missing traces; (c,d) inline and cross-line sections of cube (a), respectively; (e,f) inline and cross-line sections of cube (b), respectively.



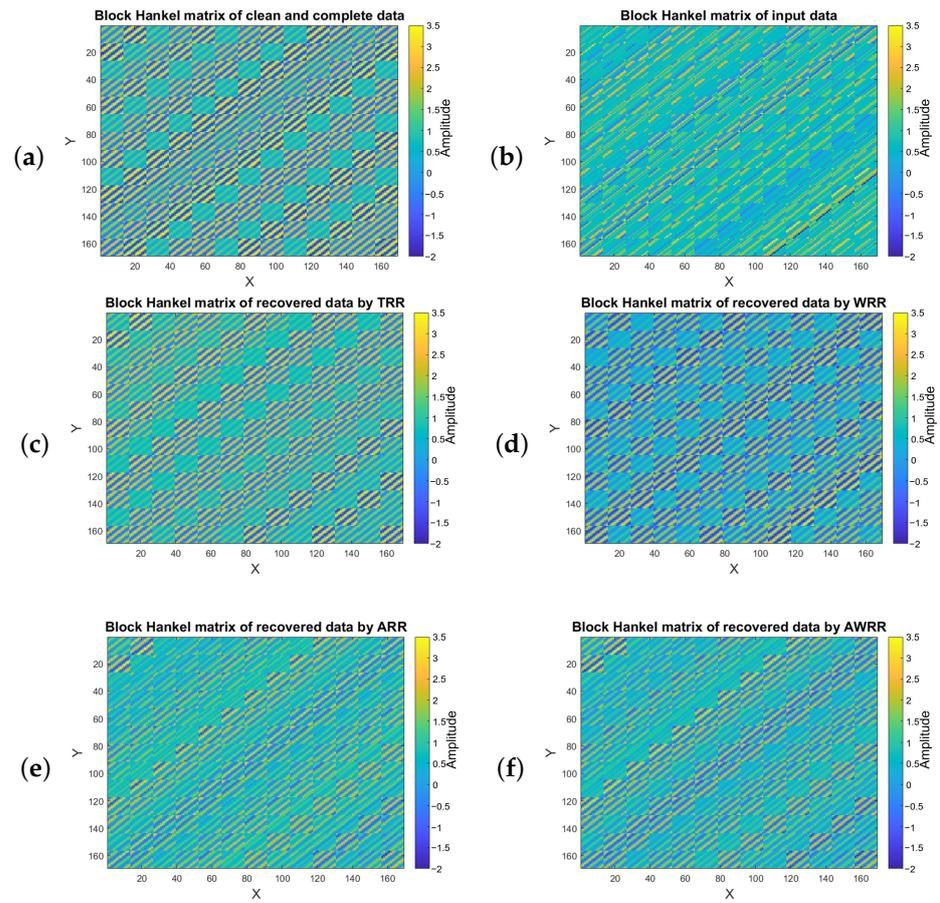
**Figure 7.** Block Hankel matrices for frequency of 20 Hz for (a) clean and complete data, (b) data with  $SNR = 2$  and 51% missing traces, (c) data recovered by TSVD rank-reduction after ten iterations, (d) data recovered by WRR after ten iterations.



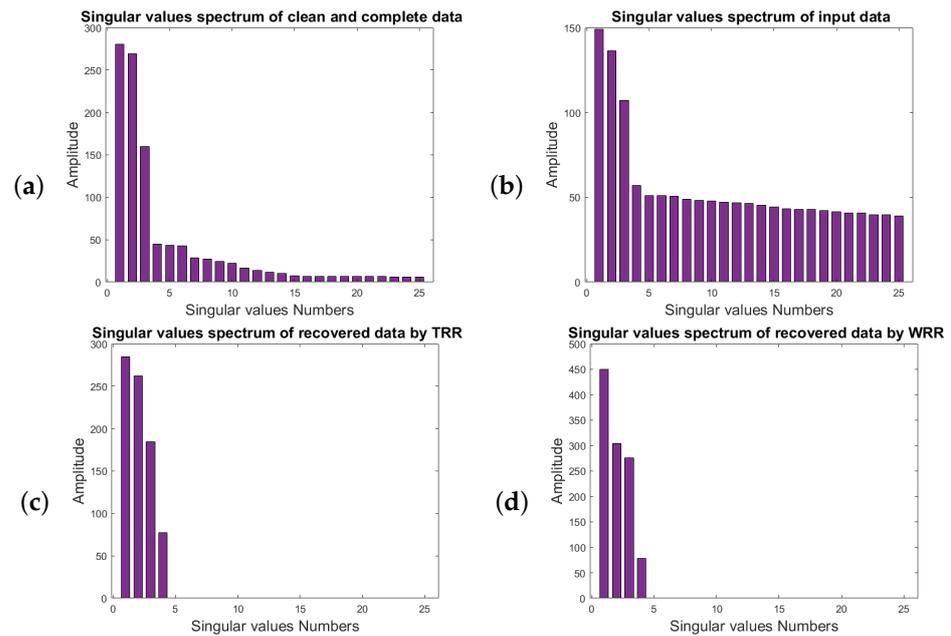
**Figure 8.** The spectrum of the first 25 singular values of the block Hankel matrices for frequency of 20 Hz for (a) clean and complete data and (b) data with  $SNR = 2$  and 51% missing traces; (c) the first 12 singular values of the data; (d) singular-value spectrum of the data after applying the weighting operator.

Figure 9 corresponds to the block Hankel matrices of the same data in the constant frequency slice of 50 Hz. Figure 9a shows the block Hankel matrix of the clean and complete data, and Figure 9b is the block Hankel matrix of data with  $SNR = 2$  where 51% of traces are removed. Figure 9c displays the recovered block Hankel matrix after applying TRR and rank = 4 for ten iterations. Figure 9d indicates the block Hankel matrix after applying the WRR with predefined rank = 4. Figure 9e represents the recovered block Hankel matrix of the ARR method, and Figure 9f is the block Hankel matrix recovered by the weighting ARR (AWRR) method. Generally, the results of AWRR and WRR are smoother than the output of TRR and ARR, and the results of ARR and AWRR include more details than those of TRR and WRR.

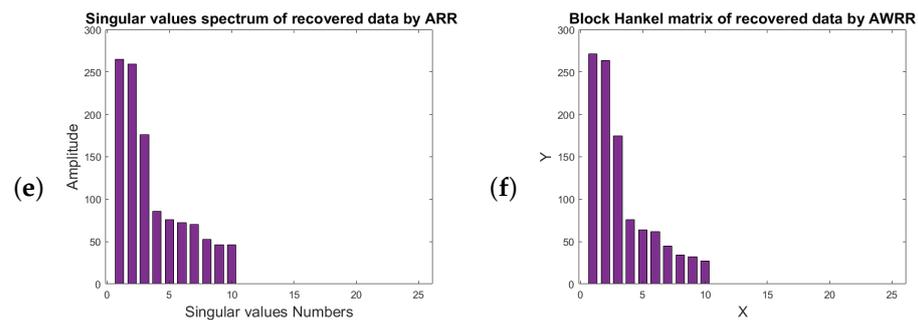
Figure 10 corresponds to the spectrum of the first 25 singular values of the block Hankel matrices of Figure 9. Figure 10a represents the singular-value spectrum of the block Hankel matrix of clean and complete data for the frequency of 50 Hz. The abrupt drop in energy of the third and the fourth singular values relates to the useful signal. However, there are still nonzero singular values that are related to the useful signal. Figure 10b is the singular-value spectrum of the block Hankel matrix of noisy and incomplete data. Figure 10c shows the singular-value spectrum of the noisy and incomplete data after TRR with rank = 4. Figure 10d displays the singular-value spectrum of the noisy and incomplete data after applying WRR with rank = 4. Figure 10e depicts the singular-value spectrum after applying ARR. Figure 10d indicates the singular-value spectrum of the noisy and incomplete data after the implementation of the AWRR method. From the figures, one can see that the result of AWRR is more comparable to the corresponding result for the desired signal in Figure 10a.



**Figure 9.** Block Hankel matrices for frequency of 50 Hz for (a) clean and complete data and (b) data with  $SNR = 2$  and 50% missing traces; (c) block Hankel matrix recovered by TRR rank = 4 after ten iterations; (d) data recovered by WRR after ten iterations. (e) Data recovered by applying the ARR method for ten iterations, and (f) data recovered after applying AWRR for ten iterations.



**Figure 10.** Cont.



**Figure 10.** The spectrum of the first 25 singular values of the block Hankel matrices for frequency slice of 50 Hz for (a) clean and complete data, (b) data with  $SNR = 2$  and 50% killed traces, (c) data recovered by TRR, (d) data recovered by WRR, (e) data recovered by ARR, (f) data recovered by AWRR.

### 3. Results and Discussion

Several evaluations are presented in this section to estimate the proficiency of the different methods of rank reduction for the MSSA algorithm. First, the methods were tested on a synthetic shot gather containing nine hyperbolic events with different curvatures. Then, they were tested on a shot gathered from a 3-D field dataset. To judge the reconstruction results numerically, we use an interpolation quality factor ( $QF$ ) defined by:

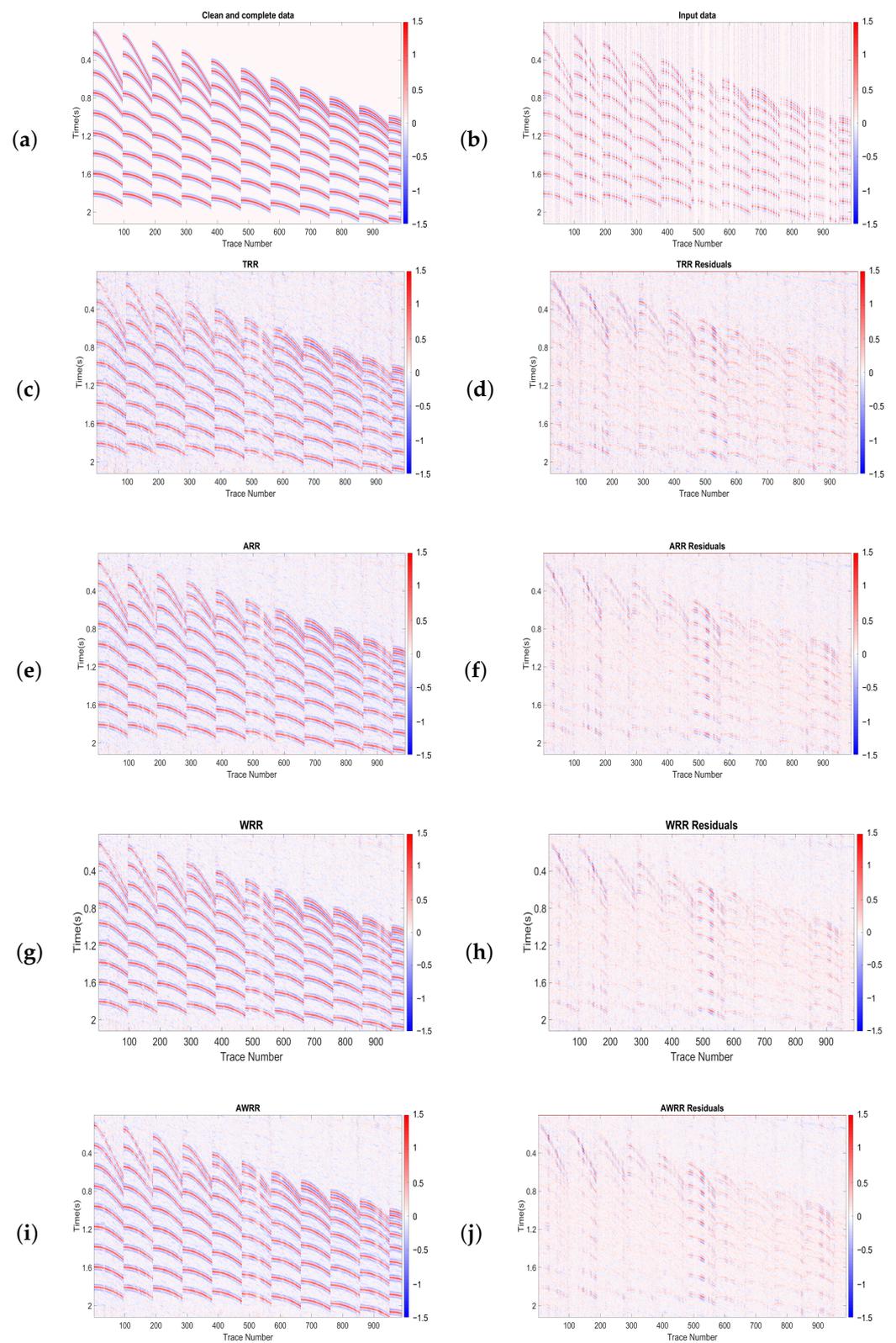
$$QF = 10 \log_{10} \left( \frac{\|d_0\|_2^2}{\|d_f - d_0\|_2^2} \right), \quad (17)$$

where  $d_0$  is the clean and complete data, and  $d_f$  is the result after applying an interpolation algorithm.

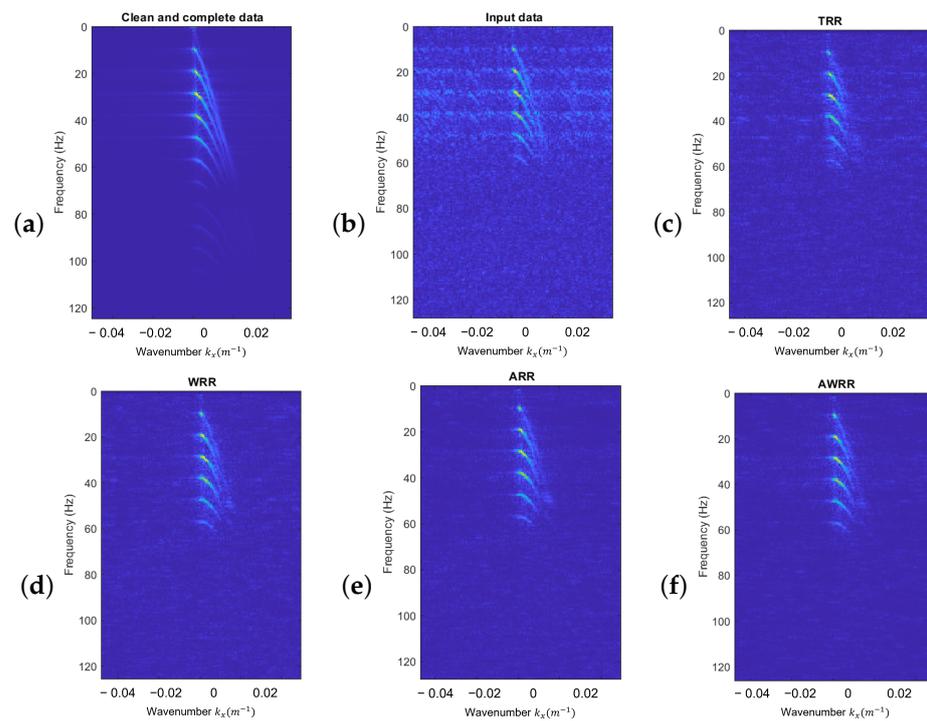
#### 3.1. Synthetic Data

Figure 11 shows the result of applying the previously discussed methods of rank-reduction to a synthetic shot gather holding nine hyperbolic events with different curvatures. The first test is a cube of 100 inline and 11 crosslines with  $SNR = 2$  and 60% decimated traces. We chose the local window with 23 traces in the inline direction and 11 in the crossline direction for each method and half a window overlapping in each direction. We set the number of iterations to be constant for all of the methods and the frequency range for the application of the MSSA algorithm to 1 to 100 Hz. The rank of TRR and WRR was set to nine. Figure 11a is the desired data arranged into a 2-D matrix. Figure 11b shows input data with  $SNR = 2$  and 60% missing traces. The results of applying TRR, ARR, WRR, and AWRR methods are shown in Figure 11c,e,g,i, respectively. Figure 11d,f,h,j are residual errors of Figure 11c,e,g,i, respectively. All four methods recovered the signals with correct amplitudes. However, in the TRR and ARR results, we can see significant residual errors. The AWRR result is much cleaner than those of TRR, ARR, and WRR. The input data quality factor is  $QF = -1.44$  dB. The output quality factors of the TRR, ARR, WRR, and AWRR methods were  $QF = 5.23, 6.58, 7.56, 8.12$  dBs, respectively.

Figure 12 compares the  $f - k$  spectra of the discussed rank-reduction methods. Figure 12a shows the  $f - k$  spectra of clean data. Figure 12b represents the  $f - k$  spectra of data with  $SNR = 2$  and 50% missing traces. Figure 12c–f are the  $f - k$  spectra of the TRR, WRR, ARR, and AWRR methods, respectively.

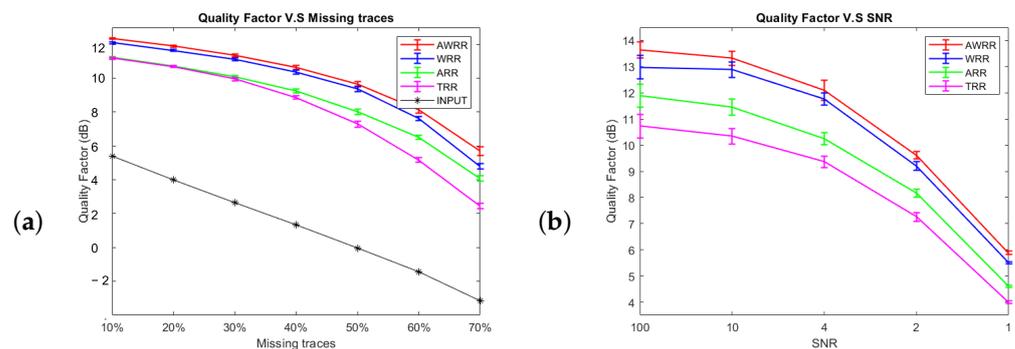


**Figure 11.** Comparison of different methods of rank-reduction. (a) Clean and complete data, (b) input data with  $SNR = 2$  and 60% missing traces; (c,e,g,i) data interpolated by TRR, ARR, WRR, and AWRR, respectively; (d,f,h,j) residual errors of TRR, ARR, WRR, and AWRR, respectively.



**Figure 12.**  $f - k$  spectral comparison of (a) clean and complete data, (b) input data, (c) TRR result, (d) WRR result, (e) ARR result, (f) AWRR result.

To evaluate the stability of each algorithm, we ran them with different percentages of gaps and noise realizations for an SNR of two. The other parameters were constant for each run. Figure 13a shows a graph of the interpolation quality factor for each method versus the gap ratio. The length of the error bars is the standard deviation of the interpolation quality factor, and each colored line connects the mean value of the interpolation quality factor for each method at each gap ratio. As we can see in Figure 13a, the input quality factor decreases with increasing gap ratio, and the output quality factor decreases with increasing gap ratios. Moreover, the output quality factor curve of the AWRR method is always above the other curves.



**Figure 13.** (a) The mean and standard error of the quality of reconstruction versus gap ratio for various methods. These results were obtained by running each method on 20 noise realizations of the dataset with different gap ratios and rank = 9 for the methods requiring a predefined rank. (b) The mean and standard error of the reconstruction quality versus SNR for various methods. These results were obtained by running each algorithm on 20 noise realizations of the dataset with a gap ratio = 50%, and rank = 9 for the methods requiring a predefined rank.

In the next experiment, the stability analysis was repeated to test the sensitivity of each method to the additive noise. The test was performed by changing the level of signal-

to-noise ratio with 20 different realizations of SNR for a gap ratio = 50%. We set the other parameters to be constant for each run. In Figure 13b we can see that the proposed method (AWRR) outperforms the other methods even with poor quality of the signal.

Table 1 indicates the values of the mean and standard deviation of the interpolation quality of each method for each gap percent. We can see that the average value of the AWRR quality factor is higher than for the other methods. Regarding the discussed methods, AWRR has lower standard deviation values, which indicates the values tend to be close to the mean value, especially when the data are sparser.

**Table 1.** Mean and standard deviation of the quality of interpolation for each method for different gap-ratio inputs.

Gap Ratio	Output Quality Factor (dB)							
	AWRR		WRR		ARR		TRR	
	MEAN (dB)	STD	MEAN (dB)	STD	MEAN (dB)	STD	MEAN (dB)	STD
10%	12.34	0.04	12.10	0.04	11.22	0.06	11.17	0.04
20%	11.90	0.05	11.64	0.05	10.71	0.09	10.69	0.06
30%	11.36	0.05	11.12	0.05	10.08	0.08	8.86	0.08
40%	10.45	0.07	10.35	0.07	9.24	0.10	8.90	0.11
50%	9.64	0.16	9.38	0.20	8.01	0.17	7.28	0.20
60%	8.13	0.13	7.61	0.13	6.51	0.14	5.16	0.20
70%	5.70	0.16	4.78	0.21	4.10	0.21	2.46	0.25

Table 2 represents the values of the mean and standard deviation of the output quality factor of each method for each signal-to-noise ratio input.

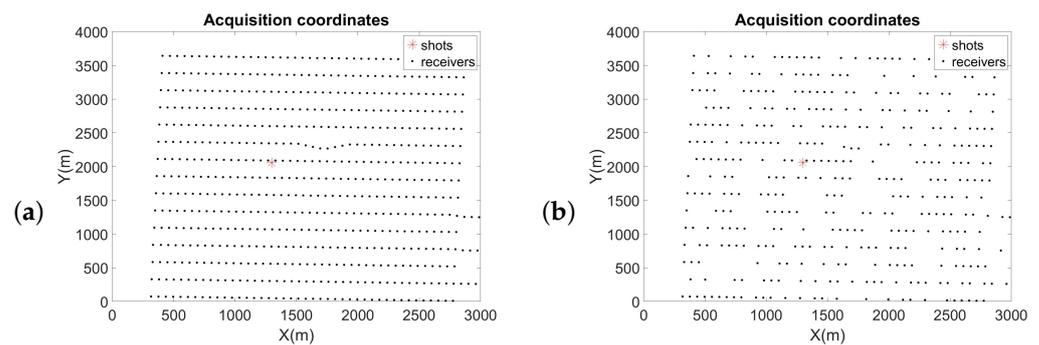
**Table 2.** Mean and standard deviation of the quality of interpolation for each method for different SNR inputs.

Input SNR	Output Quality Factor (dB)							
	AWRR		WRR		ARR		TRR	
	MEAN (dB)	STD	MEAN (dB)	STD	MEAN (dB)	STD	MEAN (dB)	STD
100	13.64	0.41	12.98	0.41	11.90	0.44	10.72	0.30
10	13.32	0.42	12.89	0.42	11.44	0.30	10.34	0.28
4	12.10	0.26	11.77	0.28	10.25	0.23	9.36	0.38
2	9.61	0.26	9.20	0.26	8.16	0.16	7.26	0.15
1	5.88	0.09	5.51	0.09	4.61	0.05	3.40	0.07

### 3.2. Real Field Data

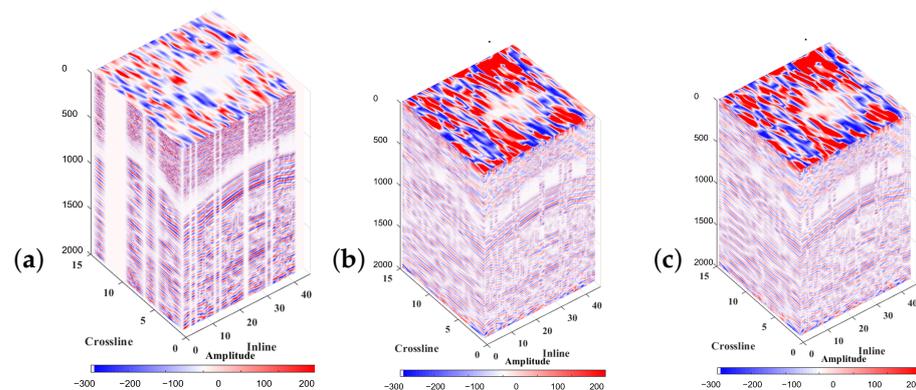
This experiment tested the efficiency of the methods on a shot gathered of a 3-D field dataset. It is easy to inspect traces in shot/receiver gather displays for poor receivers or any bad shots, which are the logistic constraints during the seismic survey. Regarding the input data, the best results are obtained with NMO-corrected data [6]. However, in this experiment, we applied the MSSA interpolation before NMO correction to see the effect of the algorithm on the curvature.

Figure 14 refers to the acquisition coordinates of the shot gather. The red star represents the location of the shot, and the black dot indicates the receiver's location. Figure 14a shows the initial distribution of the traces in a shot gather. For this experiment, 41% of the traces were removed. Figure 14b illustrates the geometry of input traces after removed 41% of them.

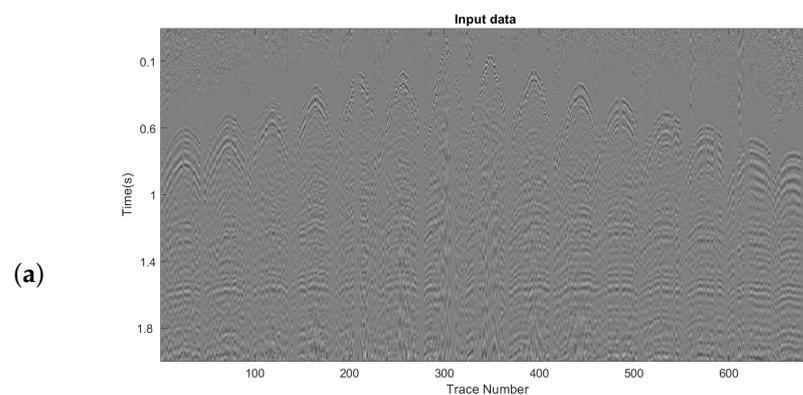


**Figure 14.** Geometry of the field data. (a) Initial distribution of the traces. (b) Distribution of traces after removing 41% of them.

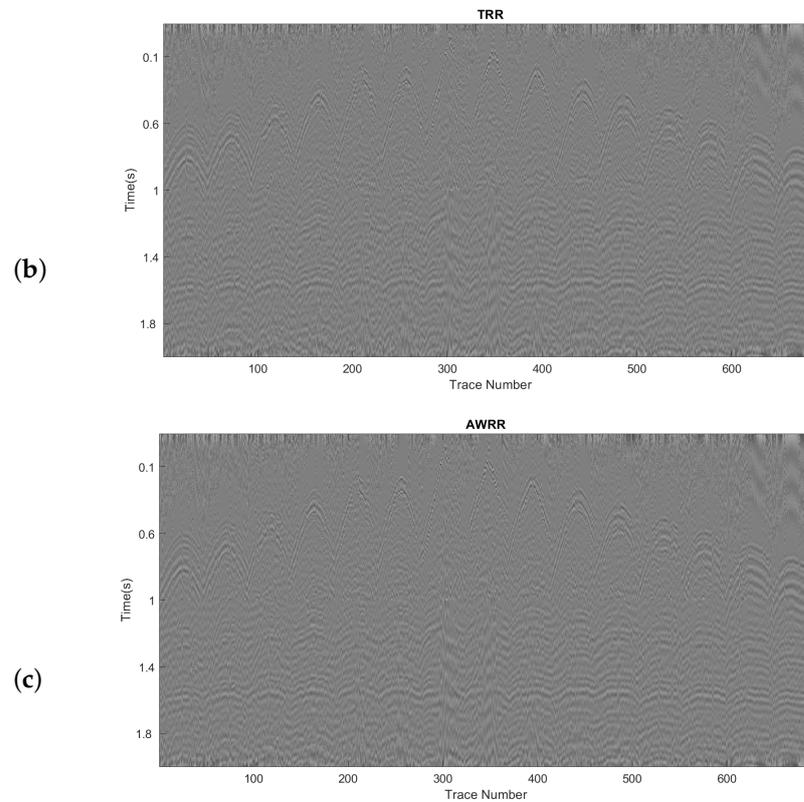
This test applied the MSSA algorithm using the AWRR and TRR methods in the rank-reduction step. Figure 15a shows the input cube of data, which presents the missing traces. Figure 15b shows the result of the interpolation using TRR in the rank-reduction step, and Figure 15c shows the result of using AWRR. For both tests, the number of iterations and processing window remained unchanged. For the TRR approach, the predefined rank was set to rank = 10. Figure 16 represents the cube of data arranged into a 2-D matrix. Figure 16a shows the input data. Figure 16b shows the result of applying TRR. Figure 16c shows the result of applying AWRR. Figure 17 corresponds to a patch of data from the time 1.35 (s) to 1.75 (s) and the trace numbers 160 to 190. Figure 17a shows the input data. Figure 17b shows the result of applying TRR. Figure 17c shows the result of applying AWRR.



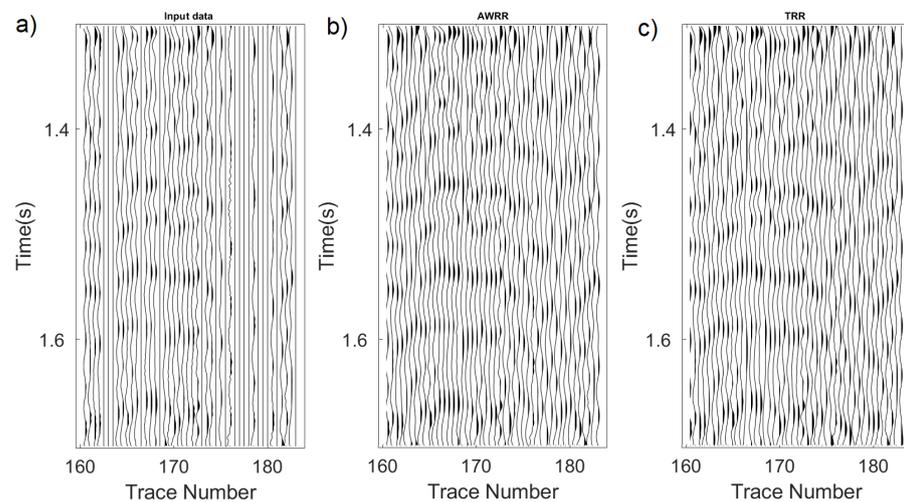
**Figure 15.** Cube of 3-D field data. (a) The input data. (b) Interpolation result of applying TRR. (c) Interpolation result of applying AWRR.



**Figure 16.** Cont.



**Figure 16.** (a) Input real data, (b) interpolation result of applying TRR approach, (c) interpolation result of applying the AWRR method.



**Figure 17.** Zooming area of Figure 16 from time 1.35 (s) to 1.75 (s) and trace numbers 160 to 190. (a) input data, (b) TRR result, (c) AWRR result.

#### 4. Conclusions

We proposed an adaptive weighting rank-reduction (AWRR) method by exploring the singular values of the Hankel matrix in each frequency slice. The method chooses the maximum energy ratio between the two following singular values as the criterion to define the optimal rank. For 3-D data, the block Hankel matrices in higher frequencies are recovered with higher ranks. Consequently, the method selects the second cutoff in the singular-value spectrum. AWRR is based on an adaptive rank-reduction method combined with a cascade weighting operator. It is adaptive in selecting the rank of the block Hankel

matrices. In addition, the weighting operator reduces the effect of additive random-noise projection on data, so a smoother block of Hankel matrices is its result. Furthermore, for different types of data, such as linear and curved, through a lot of simulations with different SNR and gap percentages, the proposed algorithm showed a higher output SNR and a better reconstruction of data with the predefined rank reduction methods. Most interpolation rank-reduction methods, such as MSSA, are SVD-based methods in their rank-reduction step. An SVD decomposition of a Hankel matrix is a very time-consuming operation. In addition, the iterative approach in the interpolation part of MSSA makes the algorithm time-consuming. To make AWRR in the MSSA algorithm applicable for higher dimensions, it is recommended to move from the SVD-based method to the SVD-free-based rank minimization method.

**Author Contributions:** Conceptualization, F.B. and D.T.; methodology, Bayati; software, Bayati and Trad; validation, F.B.; formal analysis, F.B.; investigation, F.B. and D.T.; resources, F.B. and D.T.; data curation, F.B.; writing—original draft preparation, F.B.; writing—review and editing, F.B. and D.T.; visualization, F.B.; supervision, Trad.; project administration, Trad.; funding acquisition, Trad. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19.

**Acknowledgments:** We thank the sponsors of CREWES for their continued support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Abma, R.; Claerbout, J. Lateral prediction for noise attenuation by tx and fx techniques. *Geophysics* **1995**, *60*, 1887–1896. [[CrossRef](#)]
2. Spitz, S. Seismic trace interpolation in the FX domain. *Geophysics* **1991**, *56*, 785–794. [[CrossRef](#)]
3. Porsani, M.J. Seismic trace interpolation using half-step prediction filters. *Geophysics* **1999**, *64*, 1461–1467. [[CrossRef](#)]
4. Sacchi, M.D.; Ulrych, T.J.; Walker, C.J. Interpolation and extrapolation using a high-resolution discrete Fourier transform. *IEEE Trans. Signal Process.* **1998**, *46*, 31–38. [[CrossRef](#)]
5. Liu, B.; Sacchi, M.D. Minimum weighted norm interpolation of seismic records. *Geophysics* **2004**, *69*, 1560–1568. [[CrossRef](#)]
6. Trad, D. Five-dimensional interpolation: Recovering from acquisition constraints. *Geophysics* **2009**, *74*, V123–V132. [[CrossRef](#)]
7. Sacchi, M.D.; Ulrych, T.J. Model re-weighted least-squares Radon operators. In *SEG Technical Program Expanded Abstracts 1995*; Society of Exploration Geophysicists: Houston, TX, USA, 1995; pp. 616–618.
8. Trad, D.O.; Ulrych, T.J.; Sacchi, M.D. Accurate interpolation with high-resolution time-variant Radon transforms. *Geophysics* **2002**, *67*, 644–656. [[CrossRef](#)]
9. Herrmann, F.J. Curvelet-domain matched filtering. In Proceedings of the 2008 SEG Annual Meeting, Las Vegas, NV, USA, 9–14 November 2008.
10. Naghizadeh, M.; Sacchi, M.D. Robust reconstruction of aliased data using autoregressive spectral estimates. *Geophys. Prospect.* **2010**, *58*, 1049–1062. [[CrossRef](#)]
11. Vassallo, M.; Özbek, A.; Özdemir, K.; Eggenberger, K. Crossline wavefield reconstruction from multicomponent streamer data: Part 1—Multichannel interpolation by matching pursuit (MIMAP) using pressure and its crossline gradient. *Geophysics* **2010**, *75*, WB53–WB67. [[CrossRef](#)]
12. Ghaderpour, E. Multichannel antileakage least-squares spectral analysis for seismic data regularization beyond aliasing. *Acta Geophys.* **2019**, *67*, 1349–1363. [[CrossRef](#)]
13. Stolt, R.H. Seismic data mapping and reconstruction. *Geophysics* **2002**, *67*, 890–908. [[CrossRef](#)]
14. Fomel, S. Seismic reflection data interpolation with differential offset and shot continuation. *Geophysics* **2003**, *68*, 733–744. [[CrossRef](#)]
15. Wang, B.; Zhang, N.; Lu, W.; Zhang, P.; Geng, J. Seismic data interpolation using deep learning based residual networks. In Proceedings of the 80th EAGE Conference and Exhibition 2018, Copenhagen, Denmark, 11–14 June 2018.
16. Siahkoochi, A.; Kumar, R.; Herrmann, F. Seismic data reconstruction with generative adversarial networks. In Proceedings of the 80th EAGE Conference and Exhibition 2018, Copenhagen, Denmark, 11–14 June 2018.
17. Mandelli, S.; Borra, F.; Lipari, V.; Bestagini, P.; Sarti, A.; Tubaro, S. Seismic data interpolation through convolutional autoencoder. In *SEG Technical Program Expanded Abstracts 2018*; Society of Exploration Geophysicists: Houston, TX, USA, 2018; pp. 4101–4105.
18. Mikhailiuk, A.; Faul, A. Deep learning applied to seismic data interpolation. In Proceedings of the 80th EAGE Conference and Exhibition 2018, Copenhagen, Denmark, 11–14 June 2018.
19. Wang, B.; Zhang, N.; Lu, W.; Wang, J. Deep-learning-based seismic data interpolation: A preliminary result. *Geophysics* **2019**, *84*, V11–V20. [[CrossRef](#)]

20. Trickett, S. F-xy Cadzow noise suppression. In *SEG Technical Program Expanded Abstracts 2008*; Society of Exploration Geophysicists: Houston, TX, USA, 2008; pp. 2586–2590.
21. Sacchi, M. Fx singular spectrum analysis: CSPG CSEG CWLS Convention. In *Proceedings of the Frontiers + Innovation—2009 CSPG CSEG CWLS Convention*, Calgary, AB, Canada, 4–8 May 2009; pp. 392–395.
22. Trickett, S.; Burroughs, L. Prestack rank-reducing noise suppression: Theory. In *SEG Technical Program Expanded Abstracts 2009*; Society of Exploration Geophysicists: Houston, TX, USA, 2009; pp. 3332–3336.
23. Oropenza, V.; Sacchi, M. Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis. *Geophysics* **2011**, *76*, V25–V32. [[CrossRef](#)]
24. Bayati, F.; Siahkoobi, H.R. Interpolating and denoising of seismic data by randomized SVD. In *Proceedings of the Istanbul 2012-International Geophysical Conference and Oil & Gas Exhibition*, Istanbul, Turkey, 17–19 September 2012; Society of Exploration Geophysicists and The Chamber of Geophysical: Houston, TX, USA, 2012; pp. 1–4.
25. Gao, J.; Sacchi, M.D.; Chen, X. A fast reduced-rank interpolation method for prestack seismic volumes that depend on four spatial dimensions. *Geophysics* **2013**, *78*, V21–V30. [[CrossRef](#)]
26. Naghizadeh, M.; Sacchi, M. Multidimensional de-aliased Cadzow reconstruction of seismic records. *Geophysics* **2013**, *78*, A1–A5. [[CrossRef](#)]
27. Kreimer, N.; Stanton, A.; Sacchi, M.D. Tensor completion based on nuclear norm minimization for 5D seismic data reconstruction. *Geophysics* **2013**, *78*, V273–V284. [[CrossRef](#)]
28. Kreimer, N.; Sacchi, M.D. A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation. *Geophysics* **2012**, *77*, V113–V122. [[CrossRef](#)]
29. Ely, G.; Aeron, S.; Hao, N.; Kilmer, M.E. 5D seismic data completion and denoising using a novel class of tensor decompositions. *Geophysics* **2015**, *80*, V83–V95. [[CrossRef](#)]
30. Kumar, R.; Da Silva, C.; Akalin, O.; Aravkin, A.Y.; Mansour, H.; Recht, B.; Herrmann, F.J. Efficient matrix completion for seismic data reconstruction. *Geophysics* **2015**, *80*, V97–V114. [[CrossRef](#)]
31. Rekapalli, R.; Tiwari, R.; Sen, M.K.; Vedanti, N. 3D seismic data de-noising and reconstruction using multichannel time slice singular spectrum analysis. *J. Appl. Geophys.* **2017**, *140*, 145–153. [[CrossRef](#)]
32. Siahsar, M.A.N.; Gholtashi, S.; Torshizi, E.O.; Chen, W.; Chen, Y. Simultaneous denoising and interpolation of 3-D seismic data via damped data-driven optimal singular value shrinkage. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1086–1090. [[CrossRef](#)]
33. Carozzi, F.; Sacchi, M.D. Interpolated multichannel singular spectrum analysis: A reconstruction method that honors true trace coordinates. *Geophysics* **2021**, *86*, V55–V70. [[CrossRef](#)]
34. Nadakuditi, R.R. Optshrink: An algorithm for improved low-rank signal matrix denoising by optimal, data-driven singular value shrinkage. *IEEE Trans. Inf. Theory* **2014**, *60*, 3002–3018. [[CrossRef](#)]
35. Chen, Y.; Zhang, D.; Jin, Z.; Chen, X.; Zu, S.; Huang, W.; Gan, S. Simultaneous denoising and reconstruction of 5-D seismic data via damped rank-reduction method. *Geophys. J. Int.* **2016**, *206*, 1695–1717. [[CrossRef](#)]
36. Wu, J.; Bai, M. Adaptive rank-reduction method for seismic data reconstruction. *J. Geophys. Eng.* **2018**, *15*, 1688–1703. [[CrossRef](#)]
37. Benaych-Georges, F. Rectangular random matrices, related convolution. *Probab. Theory Relat. Fields* **2009**, *144*, 471–515. [[CrossRef](#)]
38. Benaych-Georges, F.; Nadakuditi, R.R. The singular values and vectors of low rank perturbations of large rectangular random matrices. *J. Multivar. Anal.* **2012**, *111*, 120–135. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.