RANK-CONSTRAINED SEISMIC DATA INTERPOLATION AND DENOISING

Quézia Cavalcante[1*] and Milton J. Porsani[2,3]

[1]Petróleo Brasileiro SA., Rio de Janeiro, RJ, Brazil
[2]Universidade Federal da Bahia, Centro de Pesquisa em Geofísica e Geologia, Salvador, BA, Brazil
[3]Instituto Nacional de Ciência e Tecnologia de Geofísica do Petróleo, Salvador, BA, Brazil

*corresponding author email: queziacs@yahoo.com.br

ABSTRACT. Rank-constrained seismic data interpolation methods have been used to cope with spatial sampling deficiencies, but some fundamental aspects are often neglected. Understanding their underlying features is the first step for developing new solutions to overcome existing limitations. We intend to provide an intuitive description regarding low-rank strategies using the similarities between irregular samplings and noise in terms of their eigenimage representation. The interpretation of data recovery as iterative denoising helps to clarify how the traces are retrieved and the role of the rank. To emphasize either signal recovery or denoising along with the iterations, we explore non-linear versions of the decreasing weighting factor that drives the reinsertion of original samples. This type of weighting factor shows superior denoising results when raised to an integer power. Simple synthetic numerical examples illustrate the mechanics of low-rank procedures and their responses to different parameters. We also show 3D field data examples from a land survey to demonstrate the robustness of reduced-rank approaches.

Keywords: low rank, seismic data reconstruction, noise attenuation, robust seismic data interpolation

INTRODUCTION

Seismic data are supposed to be acquired in the field u s ing a u n iform s a mpling a n d f o llowing a pre-defined survey design, where the sources and receivers are distributed according to a template (Vermeer, 2012). The traces are sampled in time and space at constant rates, which define t he r espective limiting Nyquist frequencies, as stated by Shannon's theorem (Xu et al., 2005). The energy beyond the Nyquist frequency is aliased, appearing at wrong places in the Fourier spectra. Adequate time sampling is a well-resolved task, whereas the spatial directions may suffer from limitations related to acquisition costs. Additionally, obstacles and other environmental issues often lead to nonuniform spatial samplings. Some reorganizations of the traces required by processing tools may also generate spatial irregularities.

Diverse data recovery strategies have been proposed to overcome those sampling deficiencies: regular dealiasing interpolators using local dips (Abma and Kabir, 2005) or prediction filters (Spitz, 1991; Porsani, 1999), which have already been adapted to deal with irregular data (Naghizadeh and Sacchi, 2010a; Liu and Fomel, 2011); methods using Radon (Trad, 2003), *curvelet* (Naghizadeh and Sacchi, 2010b; Zhang et al., 2020), and Fourier transforms (Duijndam et al., 1999; Liu and Sacchi, 2004; Xu et al., 2005; Abma and Kabir, 2006; Zwartjes and Sacchi, 2007; Trad, 2009; Naghizadeh, 2012; Qin et al., 2018); wave-equation approaches (Stolt, 2002; Fomel, 2003; Kaplan et al., 2010); and machine-learning techniques (Jia and Ma, 2017; Wang et al., 2019; Kaur et al., 2021).

Low-rank methods constitute another important category, which has been largely investigated in recent years (Trickett et al., 2010; Oropeza and Sacchi, 2011; Kreimer and Sacchi, 2012; Ely et al., 2015; Carozzi and Sacchi, 2021; Oboué et al., 2021; Cavalcante and Porsani, 2022). They assume that regular seismic data can be represented as low-rank matrices or tensors. Thus, the aforementioned irregularities increase their rank and this fact can be used to recover the missing samples by a reduced-rank approximation. Multidimensional data may be expressed directly as a tensor (Kreimer and Sacchi, 2012) or embedded into a Hankel or a Toeplitz matrix to perform the rank reduction. The last procedure is known as Cadzow filtering or multichannel singular spectrum analysis (MSSA) (Oropeza and Sacchi, 2011; Gao et al., 2013). Although a coarse, regular sampling may not affect this type of representation, it still can be used for

dealiasing, after some adaptation (Naghizadeh and Sacchi, 2013; Huang et al., 2020). It is also possible to treat multidimensional signals looking through one spatial direction at a time, as an extension of the eigenimage filter (Trickett and Burroughs, 2009; Cavalcante and Porsani, 2021). In some sense, this is equivalent to unfolding the data tensor, a widely-used operation in multilinear algebraic approaches (Kolda and Bader, 2009; Kreimer et al., 2013; Carozzi and Sacchi, 2019).

Here we intend to provide a meaningful description of the rank-constrained interpolation methods. We explore the similarities between irregular samplings and uncorrelated noise in terms of their eigenimage representation (Freire and Ulrych, 1988). Instead of thinking of denoising as a welcome side-effect of low-rank recovery (common to other techniques as well), we consider the reconstruction itself as a type of iterative denoising. We use the basic form of eigenimage reconstruction (Trickett and Burroughs, 2009; Trickett et al., 2010; Cavalcante and Porsani, 2021) to perform the analysis, but the same interpretation remains valid for low-rank recovery approaches in general.

Additionally, we evaluate the role of the rank and the weighting factor which controls the reinsertion of original samples in the iterative optimization scheme. This weight can assume at least two different forms: a constant value (Kreimer and Sacchi, 2012; Gao et al., 2013) and iterative-dependent values with either a linear or a non-linear behavior (Oropeza and Sacchi, 2011; Huang et al., 2020). We propose novel approaches for this parameter and discuss their strengths and weaknesses both qualitatively and quantitatively, using synthetic numerical experiments. We also show field data examples to better illustrate some points.

2   THEORY

2.1   Data matrix and its low-rank approximation

A complex function of three variables $D_{obs}(f, x, y)$ represents a 3D seismic volume in the frequency-space domain, where f is the temporal frequency and (x, y) are spatial coordinates. By considering irregularly spaced data at a regular grid, the function $D_{obs}(f, x, y)$ becomes a tensor $\mathcal{D}_{obs}$ of dimensions

$N_x \times N_y \times N_f$. For each constant frequency f, the seismic data may be written as a matrix $\mathbf{D}_{obs}(f)$ of dimensions $N_x \times N_y$. The tensor and matrix representations assume that multiple samples falling in the same cell are averaged, whereas the elements within empty cells are set to zero. We will drop the dependency in f from now on, but keep in mind that the analysis is done for each frequency slice independently. For now, we also suppress the subscript of $\mathbf{D}_{obs}$ to simplify the notation. Thus, consider a matrix $\mathbf{D}$, which comes from a complete and noiseless data volume.

The computation of a reduced-rank version $\widetilde{\mathbf{D}}$ of the matrix $\mathbf{D}$ (a key step of low-rank strategies) primarily uses the singular value decomposition (SVD) (Golub and Van Loan, 1996). For a given rank r, the approximation $\widetilde{\mathbf{D}}$ is the solution to an optimization problem which minimizes (Oropeza and Sacchi, 2011; Menke, 2012)

$$J = \|\mathbf{D} - \widetilde{\mathbf{D}}\|_F^2 = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} |d_{ij} - \tilde{d}_{ij}|^2, \tag{1}$$

where $\|.\|_F$ is the Frobenius norm (or matrix norm) and $|d_{ij} - \tilde{d}_{ij}|$ represents the absolute value of the error between the (i, j) element of $\mathbf{D}$ and $\widetilde{\mathbf{D}}$, respectively. Thereby, the SVD-based $\widetilde{\mathbf{D}}$ is the best approximation in a least-squares sense (Eckart and Young, 1936):

$$\widetilde{\mathbf{D}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H. \tag{2}$$

The notation $[.]^H$ symbolizes the conjugate transpose. The columns of the $N_x \times r$ matrix $\mathbf{U}$ are the first r left singular vectors of $\mathbf{D}$ (eigenvectors of the covariance matrix $\mathbf{D}\mathbf{D}^H$). The rows of the $r \times N_y$ matrix $\mathbf{V}^H$ are the first r right singular vectors of $\mathbf{D}$ (eigenvectors of the covariance matrix $\mathbf{D}^H\mathbf{D}$). The $r \times r$ matrix $\mathbf{\Sigma} = diag(\sigma_1, \ldots, \sigma_r)$ is formed by the r largest singular values of $\mathbf{D}$ in decreasing order. They are the positive square roots of the first r eigenvalues of both covariance matrices. We may rewrite equation 2 as the sum

$$\widetilde{\mathbf{D}} = \sum_{k=1}^{r} \sigma_k \mathbf{u}_k \mathbf{v}_k{}^{H}. \tag{3}$$

The rank-1 matrix resulting from the outer product $\mathbf{u}_k \mathbf{v}_k{}^{H}$ is called the k-th *eigenimage* of $\mathbf{D}$ and $\sigma_k \mathbf{u}_k \mathbf{v}_k{}^{H}$ is known as the k-th *weighted eigenimage* (Freire and Ulrych, 1988; Trickett and Burroughs, 2009). Note that the vectors $\mathbf{u}_k$ and $\mathbf{v}_k$ are the k-th left and right singular vectors, which appear as the columns and rows of $\mathbf{U}$ and $\mathbf{V}^{H}$, respectively. Equations 2 and 3 (known as truncated SVD), may also be expressed as (Freire and Ulrych, 1988)

$$\widetilde{\mathbf{D}} = \mathbf{D}\mathbf{V}\mathbf{V}^{H} = \sum_{k=1}^{r} \mathbf{D}\mathbf{v}_k \mathbf{v}_k{}^{H} \tag{4}$$

or

$$\widetilde{\mathbf{D}} = \mathbf{U}\mathbf{U}^{H}\mathbf{D} = \sum_{k=1}^{r} \mathbf{u}_k \mathbf{u}_k{}^{H}\mathbf{D}. \tag{5}$$

It can be shown that $\mathbf{D}$ has a rank that corresponds to the number of linear events contained in the seismic volume (Trickett and Burroughs, 2009). Because amplitude variation along the events may increase the rank of the data matrix, preprocessing including amplitude equalization is a good practice. Low-rank approaches work with curved events as well, as long as they can be regarded as a superposition of linear events. Windowed processing may be used to ensure that premise, which is also common to Fourier-based methods (Abma and Kabir, 2006).

It is worth mentioning that there is a direct relationship between the rank-r approximation $\widetilde{\mathbf{D}}$, the r principal components, and the Karhunen-Loève (KL) transform of $\mathbf{D}$, which is given by (Freire and Ulrych, 1988; Ulrych et al., 1999)

$$\mathbf{P} = \mathbf{U}^{H}\mathbf{D} = \mathbf{\Sigma}\mathbf{V}^{H}. \tag{6}$$

The inverse KL transform is

$$\widetilde{\mathbf{D}} = \mathbf{U}\mathbf{P}, \tag{7}$$

which is equivalent to equation 5.

According to equation 4, $\widetilde{\mathbf{D}}$ can be calculated using the eigenvectors of the $N_y \times N_y$ covariance matrix $\mathbf{D}^H\mathbf{D}$, obtained by the eigenvalue decomposition (EVD)

$$\mathbf{D}^H\mathbf{D} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H, \tag{8}$$

where $\mathbf{\Lambda} = diag(\lambda_1, \ldots, \lambda_{N_y})$ and the eigenvalues $\lambda_j = \sigma_j{}^2$. Note that equation 8 is correct only if the matrix $\mathbf{V}$ contains the whole set of eigenvectors. Otherwise, it would represent a low-rank approximation of the covariance matrix itself. There is no loss of generality if $N_y > N_x$, which means that $\mathbf{D}$ has a maximum of N_x singular values and the remaining λ_j, $j = N_x, \ldots, N_y$ are equal to zero. The subsequent discussion could use the other covariance matrix as well, but we turn our attention to $\mathbf{D}^H\mathbf{D}$ just to stay in line with the theory discribed by Cavalcante and Porsani (2021). Instead of analyzing the singular values σ_j, we will look at the eigenvalues λ_j. This is reasonable given that the percentage of energy contained in the reduced-rank approximation is given by (Freire and Ulrych, 1988)

$$E(r) = \frac{\sum_{j=1}^{r}\sigma_j{}^2}{\sum_{j=1}^{N_y}\sigma_j{}^2} = \frac{\sum_{j=1}^{r}\lambda_j}{\sum_{j=1}^{N_y}\lambda_j}. \tag{9}$$

The denominators in equation 9, which give the total energy, are also related to the Frobenius norm of $\mathbf{D}$ and the trace of the covariance matrices, such that $\|\mathbf{D}\|_F^2 = \sum_{j=1}^{N_y}\sigma_j{}^2 = \mathrm{trace}(\mathbf{D}^H\mathbf{D}) = \mathrm{trace}(\mathbf{D}\mathbf{D}^H)$ (Ford, 2014).

Recall that we are looking at individual frequency slices, in the frequency-space domain. Thus, to retrieve the 3D seismic volume one should apply the procedure for all frequencies and compute the inverse Fourier transform. This operation results in a *weighted eigenvolume* if it is performed using only one of the weighted eigenimages (see equation 3). Figure 1 illustrates a 3D synthetic volume with linear events and its first and second weighted eigenvolumes. Their dimensions are $64 \times 64 \times 204$, referring to N_x, N_y, and N_t (number of time samples). The first eigenvolume captures the most energetic features (the two

dipping planes with the same slope along the y-direction). Figure 1d shows the rank-2 approximation as well, which is sufficient to represent all the information. Figure 2 portrays the first 16 eigenvalues of $\mathbf{D}^H\mathbf{D}$ corresponding to the frequency slice $f = 29.3$ Hz. As expected, only two of them have significant values.

2.2   Covariance matrix and noise

The $N_x \times N_y$ observed data matrix $\mathbf{D}_{obs}$ at a constant frequency, with additive noise, can be written as

$$\mathbf{D}_{obs} = \mathbf{D} + \mathbf{N}, \tag{10}$$

where $\mathbf{N}$ is an $N_x \times N_y$ matrix containing zero-mean, spatially-white random noise, independent from trace to trace. Let's say, for simplicity, that there is only one linear event with slowness equal to zero (a horizontal event). In this case, the $N_x \times N_y$ signal matrix $\mathbf{D}$ is formed by constant-valued elements (Cavalcante and Porsani, 2021).

The $N_y \times N_y$ data covariance matrix is composed of signal and noise, so that (Kirlin, 1999)

$$\begin{aligned} \mathbf{C} &= \mathbf{D}_{obs}^H \mathbf{D}_{obs} \\ &= (\mathbf{D} + \mathbf{N})^H (\mathbf{D} + \mathbf{N}) \\ &= \mathbf{D}^H\mathbf{D} + \mathbf{N}^H\mathbf{N} + \mathbf{N}^H\mathbf{D} + \mathbf{D}^H\mathbf{N} \\ &= \mathbf{D}^H\mathbf{D} + \mathbf{N}^H\mathbf{N} + 2Re\left\{\mathbf{D}^H\mathbf{N}\right\}. \end{aligned} \tag{11}$$

Because we are dealing with uncorrelated noise, the cross-terms $2Re\left\{\mathbf{D}^H\mathbf{N}\right\}$ vanish. The noise covariance matrix is given by

$$\mathbf{N}^H\mathbf{N} = \sigma_N^2\mathbf{I}, \tag{12}$$

where σ_N^2 is the variance of the noise and $\mathbf{I}$ is the $N_y \times N_y$ identity matrix. Thus, the data covariance

matrix can be written as

$$\mathbf{C} = \mathbf{D}^H \mathbf{D} + \sigma_N^2 \mathbf{I}. \tag{13}$$

The horizontal event leads to a signal covariance matrix $\mathbf{D}^H \mathbf{D} = w^2 N_y \mathbf{1}$, where $\mathbf{1}$ is an $N_y \times N_y$ matrix of ones (the identity matrix under Hadamard multiplication). The constant w is the seismic wavelet's component in frequency f. $\mathbf{D}^H \mathbf{D}$ has a rank equal to one, which means that it has only one eigenvalue λ_D. By multiplying both sides of equation 13 with $\mathbf{v}_1$, which is the first eigenvector of $\mathbf{D}^H \mathbf{D}$, we obtain

$$
\begin{aligned}
\mathbf{C}\mathbf{v}_1 &= \mathbf{D}^H \mathbf{D}\mathbf{v}_1 + \sigma_N^2 \mathbf{I}\mathbf{v}_1 \\
&= \lambda_D \mathbf{v}_1 + \sigma_N^2 \mathbf{v}_1 \\
&= \left(\lambda_D + \sigma_N^2 \right) \mathbf{v}_1.
\end{aligned}
\tag{14}
$$

Therefore, $\lambda_D + \sigma_N^2$ is the largest eigenvalue of $\mathbf{C}$. This is true even with a non-horizontal linear event, where the signal matrix $\mathbf{D}$ contains a 2D complex sinusoid (Canales and Lu, 1993). The remaining $N_y - 1$ eigenvalues of $\mathbf{C}$ are $\lambda_2 = \lambda_3 = \ldots = \lambda_{N_y} = \sigma_N^2$, which indicate a measure of how close the first eigenimage of $\mathbf{D}_{obs}$ is to the signal $\mathbf{D}$. The positive-semidefinite (and Hermitian) nature of $\mathbf{C}$ means that its eigenvalues are always real and non-negative (Freire and Ulrych, 1988; Strang, 2018).

The extension for a greater number of seismic events is straightforward. As stated before, in a more general case, $\mathbf{D}$ has a rank equal to the number r of linear events (Trickett and Burroughs, 2009). Therefore, $\mathbf{D}^H \mathbf{D}$ has r significant eigenvalues and the last $N_y - r$ are equal to zero. For matrix $\mathbf{C}$, the smallest $N_y - r$ eigenvalues are equal to σ_N^2. This fact may be used to choose the best value for r, which is not known a priori in practical applications. The deviation caused by the presence of noise described by equation 12 does not alter the eigenvectors so that they are the same eigenvectors of the signal's covariance matrix $\mathbf{D}^H \mathbf{D}$ (Strang, 2018). In other words, the eigenvectors are insensitive to white noise (Ulrych et al., 1999).

The rank-r approximation of $\mathbf{D}_{obs}$ is itself an approximation of the signal matrix $\mathbf{D}$ because noise has

equal influence in the so-called "signal" and "noise" subspaces, associated with the r largest and the $N_y - r$ smallest eigenvalues, respectively. The signal is recovered indeed, whereas the noise is suppressed in a manner equivalent to an average stack (Ulrych et al., 1999).

If the noise comes from several sources of similar size, it can be described by a Gaussian (or Normal) probability density function ($p.d.f$), because it is the limiting $p.d.f$ for the sum of independent random variables, according to the *central limit theorem*. Also, linear functions of Gaussian random variables are themselves Gaussian (Menke, 2012). For these reasons, we use Gaussian additive noise in our numerical experiments, but we agree that assuming Gaussian and spatially-uncorrelated noise may represent an oversimplification in some situations (Duijndam et al., 1999).

Figure 3 shows the same 3D synthetic volume with three linear events and random noise added to obtain a signal-to-noise ratio (S/N) equal to four. The first two eigenvolumes are represented as well, along with the low-rank approximation using $r = 2$. Likewise the clean volume, $r = 2$ is a necessary and sufficient condition to retrieve the information. The noise is attenuated in the reduced-rank version, but it is still there, whereas the remaining eigenvolumes (not shown here) contains only noise. The graphic of the eigenvalues in Figure 4 reinforces this idea. It is reasonable to think that a choice of $r > 2$ would produce a noisier output.

2.3 Spatial irregularities

Consider now the observed data matrix $\mathbf{D}_{obs}$ at a constant frequency with spatial irregularities. It is related to the (ideal) complete signal matrix $\mathbf{D}$ through the sampling operator $\mathbf{S}$, such that

$$\mathbf{D}_{obs} = \mathbf{S} \odot \mathbf{D}, \tag{15}$$

where $\odot$ is the Hadamard or elementwise product and the entries of $\mathbf{S}$ are given by

$$s_{ij} = \begin{cases} 0, & \text{if the point } (i,j) \text{ is missing;} \\ \\ 1, & \text{if the point } (i,j) \text{ is not missing.} \end{cases} \tag{16}$$

If $\mathbf{S}$ contains randomly distributed elements, the rank of the resulting data matrix increases in a way that resembles the one with uncorrelated noise (Figure 4). However, there is a fundamental distinction: the signal leakage caused by spatial irregularities reduces the most significant eigenvalues when compared to the noisy case. In Figure 5, nearly 50% of the traces were randomly removed from the previous 3D clean synthetic volume. The first two eigenvolumes, however, contain both band-limited noise and signal, but with lower amplitude than the original volume (see the reduced eigenvalues in Figure 6). The remaining eigenvolumes contain only noise and the superposition of all eigenvolumes reproduces the gaps. Figure 6 also compares the eigenvalues for three decimation rates: 30%, 50%, and 70%. The leakage and the resulting signal-energy loss are more severe for a greater number of missing samples, but the noise stemming from irregularities stays at the same level.

Interestingly, the signal leakage throughout the spatial directions is responsible for filling the gaps. The random irregularities behave as noise, in some sense, but the signal is present and it is captured by the dominant weighted eigenimages. Recall that linear events correspond to bidimensional complex sinusoids in the frequency-space domain. Although the irregularities force the covariance matrix towards a diagonal one, it still measures the spatial correlation of those periodic functions.

Figure 7 portrays the 3D volume with Gaussian noise and 50% of randomly missing traces. Naturally, the eigenvalues resulting from the combination of uncorrelated noise and irregularities show a superposition of both effects. However, because the noise itself is also subject to energy leakage, its eigenvalues decrease as well, as we see in Figure 8.

Random irregularities appear as pervasive noise in the frequency-wavenumber domain as well (Naghizadeh and Sacchi, 2010c) and this fact is used by Fourier-based reconstruction methods (Abma and Kabir, 2006; Xu et al., 2005). This is closely related to the compressive sensing framework (Donoho,

2006), which comprises, basically: random sampling, sparsity in some domain, and signal recovery from just a small number of coefficients (Herrmann, 2010). If we think of the eigenimages as a basis, their respective coefficients are the singular values (Freire and Ulrych, 1988). Thus, sparsity and low rank are just two sides of the same coin.

If $\mathbf{S}$ contains a zero-valued column or row or if it represents a regular decimation (a coarse sampling, in practice), the rank of $\mathbf{D}_{obs}$ is not affected. Zero-valued columns are always linear combinations of the others. Naghizadeh and Sacchi (2013) have proposed an adaptation for the Cadzow/MSSA approaches (which represent the data as Hankel or Toeplitz matrices) that is capable of dealing with regular decimation. This is still an open field for the eigenimage filter presented here, directly using the data volumes or matrices without any *Hankelization* operation.

2.4 Data completion and denoising

In the previous section, we show how irregularities manifest themselves as noise. Therefore, recovery of missing samples and denoising are performed simultaneously, in a natural way. Once we have the reduced-rank approximation $\widetilde{\mathbf{D}}$ of the observed data matrix, our reconstruction follows the iterative strategy described by Oropeza and Sacchi (2011), where the original samples are reinserted in each iteration. This procedure guarantees that the signal energy (which suffers from the leakage illustrated in Figures 5, 6, and 7) will be retrieved. A similar approach is used by Abma and Kabir (2006) in the context of Fourier reconstruction. The recovered matrix is

$$\mathbf{D}^\nu = \mathbf{D}_{obs} + (\mathbf{1} - \mathbf{S}) \odot \widetilde{\mathbf{D}}^{\nu-1}, \tag{17}$$

where $\nu = 1, \ldots, \nu_{max}$ represents iteration index, ν_{max} is the maximum number of iterations, and $\mathbf{D}^0 = \mathbf{D}_{obs}$. The matrix $\mathbf{1}$ contains ones and has the same size as $\mathbf{S}$. Clearly, equation 17 only works for clean seismic data. Otherwise, the noisy samples would be reinserted in every iteration. To attenuate this effect,

$\mathbf{D}_{obs}$ may be replaced by the weighted average $a^{\nu}\mathbf{D}_{obs} + (1 - a^{\nu})\,\mathbf{S} \odot \widetilde{\mathbf{D}}^{\nu-1}$, so that (Oropeza and Sacchi,

2011; Kreimer and Sacchi, 2012)

$$\mathbf{D}^{\nu} = a^{\nu}\mathbf{D}_{obs} + (\mathbf{1} - a^{\nu}\mathbf{S}) \odot \widetilde{\mathbf{D}}^{\nu-1}. \tag{18}$$

The weighting factor a may assume different forms. Here we compare the performances of a constant

value $a \in (0, 1)$ (Kreimer and Sacchi, 2012; Gao et al., 2013) with an iteration-dependent value $a \in [0, 1]$

which decreases from $a^{1} = 1$ to $a^{\nu_{max}} = 0$ (Oropeza and Sacchi, 2011). In the last case, we may write

$$a = \left(\frac{\nu_{max} - \nu}{\nu_{max} - 1}\right)^{\frac{1}{p}}, \tag{19}$$

using a fractional exponent. When $p = 1$, the parameter a decreases linearly, as proposed by Oropeza

and Sacchi (2011), whereas $p > 1$ leads to non-linear behavior. Huang et al. (2020) also propose a non-

linear factor which, on the contrary, increases from 0 to 1, in a slightly different way. But their aim is similar

to ours: emphasizing signal recovery in the earlier iterations and intensifying noise removal in the final

stages. Figure 9 illustrates the weighting factor a as a function of the iteration number for different values

of p (equation 19).

Another possibility for the iteration-dependent a, using an integer exponent, is given by

$$a = \left(\frac{\nu_{max} - \nu}{\nu_{max} - 1}\right)^{p}. \tag{20}$$

Now, the earlier stages corresponding to signal recovery are sacrificed to emphasize noise removal in

the last iterations. When $p = 1$, the previous equations lead to the same behavior, where the parameter

a decreases linearly. Figure 10 displays the weighting factor a as a function of the iteration number for

different values of p (equation 20).

Gao et al. (2015) and Carozzi and Sacchi (2019) use the parallel matrix factorization for low-rank seismic tensor reconstruction, where expressions like equation 18 arise. The single-imputation algorithm described by Tomasi and Bro (2005) to cope with missing data in PARAFAC multilinear decomposition (Kolda and Bader, 2009) also leads to a similar equation.

Because the algorithm iteratively improves a solution, a stop criterion can be used to avoid unnecessary computation once the relative change (in terms of the squared Frobenius norm) between consecutive iterations is less than a small ϵ. Mathematically, we may terminate the process when

$$\frac{\|\mathbf{D}^{\nu} - \mathbf{D}^{\nu-1}\|_F^2}{\|\mathbf{D}^{\nu-1}\|_F^2} < \epsilon, \tag{21}$$

but this type of interruption might become a problem when using an iteration-dependent a. Also, care must be exercised in the choice of ν_{max}, because a fast convergence may suppress the final intended denoising effect.

3   MORE NUMERICAL EXPERIMENTS

The same toy volume with three linear events is used to demonstrate how low-rank methods work and most importantly, the way the key parameters may influence the results. To evaluate them, we compute the quality

$$Q = \frac{\|\boldsymbol{\mathcal{D}}^{true}\|_F^2}{\|\boldsymbol{\mathcal{D}}^{rec} - \boldsymbol{\mathcal{D}}^{true}\|_F^2}, \tag{22}$$

in decibels, where the tensors $\boldsymbol{\mathcal{D}}^{true}$ and $\boldsymbol{\mathcal{D}}^{rec}$ represent the original clean volume and the recovered volume, in the time domain, respectively.

We first investigate how the user-defined rank r may affect the recovery quality. Figure 11 shows the interpolation of the 3D volume with nearly 50% of missing samples, using $r = 2$ and $\nu_{max} = 20$ (the

maximum number of iterations). The resulting volume is virtually perfect, with quality $Q = 46.9$ dB. The error contains no significant energy. Clearly, steep and conflict dips are not limiting issues. On the contrary, the volume in Figure 12a obtained with $r = 5$ has an inferior quality $Q = 19.2$ dB and the error in Figure 12b is quite relevant.

In Figure 13 we plot the quality Q as a function of r to exemplify the importance of this parameter. As we have mentioned, $r = 1$ does not suffice to carry all the information. The optimum value is $r = 2$ and beyond that value, the quality decreases rapidly. It seems that SVD-based approaches lack flexibility regarding the choice of the appropriate rank: it should be the least possible value. This fact and the computational cost have motivated the search for alternatives to the SVD, such as randomized-SVD (Oropeza and Sacchi, 2011), Lanczos bidiagonalization (Gao et al., 2013), and CUR decompositions (Cavalcante and Porsani, 2022).

Figure 14 shows the reconstruction (and simultaneous denoising) of the irregular 3D volume with Gaussian noise (S/N=4). The main parameters are $r = 2$ and $\nu_{max} = 10$. We use a constant weighting factor $a = 0.5$ in this example, which leads to a noisy seismic volume. The negative quality $Q = -0.3$ dB means that the Frobenius norm of the error is superior to the norm of the clean data volume. A constant factor $a = 0.8$ is even worse, with $Q = -2.6$ dB. A last single-filtering iteration is not sufficient to generate much better results.

Examples of iteration-dependent a with a fractional exponent (equation 19) are illustrated in Figure 15, with $p = 1$ and $p = 3$. The recovery results are fairly similar, with $Q = 2.7$ dB and $Q = 2.4$ dB, respectively. Both of them are significantly superior to those using a constant a. However, as p increases, the last denoising iterations are sacrificed.

In Figure 16 we use the other form of a, with an integer power $p = 3$ (equation 20). The quality $Q = 3.3$ dB indicates a superior denoising performance, but the signal energy is not completely restored so that the error volume in Figure 16b contains signal amplitude with opposite polarity. The explanation is the following: the fast decay of a (see Figure 10) prevents the original traces to be reinserted in the very first iterations. On the other hand, more iterations are dedicated to denoising.

Figure 17 displays the recovery quality as a function of p for both non-linear procedures. More significant differences are related to a as described by equation 20, with an integer power. For the fractional exponent, the differences are quantitatively detectable but not qualitatively apparent.

Curved events, such as reflections and diffractions, can be regarded as a superposition of linear events. However, only little curvatures are acceptable to guarantee the low-rank assumption. One can also make use of windowed-processing strategies and normal moveout correction to ensure that the seismic data can be described by a small number of linear events (or, equivalently, a low-rank matrix or tensor). Because there is possibly a spatial direction with linear behavior in multidimensional data (Abma and Kabir, 2006), it might be better to compute the reduced-rank approximation along this dimension (Cavalcante and Porsani, 2021).

When using synthetic data we can clearly see and precisely quantify whatever is necessary. That is a great advantage, but sometimes it is also a source of disappointment because we tend to not tolerate any error. Of course, in the real world, we do accept a certain noise level and sometimes we even use it in our favor.

To illustrate the robustness of low-rank eigenimage reconstruction, we use 3D land data acquired according to an orthogonal geometry (Vermeer, 2012). The common-offset volumes resulting from this type of survey design have a certain pattern of missing samples, which seems especially suitable for rank-constrained interpolation. These irregularities are just consequences of the relatively straight and widely-spaced acquisition lines (Vermeer, 2012). Before data recovery, some preprocessing is required, including the application of source- and receiver-statics corrections, ground-roll attenuation, and amplitude balance.

For this particular data volume, the common-midpoint grid is $30\,\mathrm{m}\,\times\,30\,\mathrm{m}$. The tensor dimensions are $120 \times 80 \times 401$, referring to the inline, crossline, and time sample, respectively. The time sampling rate is $4\,\mathrm{ms}$ and we restrict the frequencies to the $8\,\text{-}\,90\,\mathrm{Hz}$ band. Figure 18 portrays the reconstruction of 3D common-offset volumes using $r = 3$ and $\nu_{max} = 30$, in spatial windows of size 21×21 traces. The weighting factor a has a linearly decreasing form. In Figures 18a and 18b, the irregular volume and recovery result correspond to offset bin number 3, centered at $300\,\mathrm{m}$. In Figures 18c and 18d, both volumes

correspond to offset bin number 4, centered at $420\,$m. Good interpolations result from the aforementioned set of parameters, despite a quite noisy background. Figure 19 displays time slices at $1352\,$ms of these common-offset volumes. The grayscale is the same as Figure 18, and most of the gray is related to missing samples. As expected, the shorter offset shows a higher degree of irregularities, which are not random, but sufficiently low-correlated in terms of the covariance matrix.

4   CONCLUSION

Here we unveiled some fundamental aspects hidden in the heart of rank-constrained recovery approaches. We demonstrated how the eigenimage representation can capture the signal using both energy and spatial correlation. We believe that the interpretation of trace recovery as iterative denoising helps to clarify the origin of the retrieved data. Once the appropriate rank is chosen, which may not be a simple task, the reconstructed volumes using iterative-dependent weighting factors show superior quality. Non-linear parameters with an integer exponent lead to better denoising results when compared to the linear ones. However, the signal amplitude is not fully retrieved. Not only does understanding the basic features of reduced-rank methods help us recognizing their strengths and weaknesses, but it is also essential to overcome their limitations. The often-neglected role of the rank, for instance, should be further investigated towards flexibility. Additionally, it is worth exploring less-costly alternatives to SVD, paying special attention to those not demanding the exact value of the rank (Carozzi and Sacchi, 2019; Cavalcante and Porsani, 2022).

REFERENCES

Abma, R.; Kabir, N. Comparisons of interpolation methods. The Leading Edge **2005**, 24, 984–989.

Abma, R.; Kabir, N. 3D interpolation of irregular data with a POCS algorithm. Geophysics **2006**, 71, E91–E97. 10.1190/1.2356088.

Canales, L.L.; Lu, L. 3-D trace interpolation in the *f-x-p* domain. 63rd Annual Int. Meeting, SEG, Expanded Abstracts **1993**, pp. 1174–1176. 10.1190/1.1822326.

Carozzi, F.; Sacchi, M.D. Robust tensor-completion algorithm for 5D seismic-data reconstruction. Geophysics **2019**, 84, V97–V109. 10.1190/geo2018-0109.1.

Carozzi, F.; Sacchi, M.D. Interpolated multichannel singular spectrum analysis: A reconstruction method that honors true trace coordinates. Geophysics **2021**, 86, V55–V70. 10.1190/geo2019-0806.1.

Cavalcante, Q.; Porsani, M.J. Prestack seismic data reconstruction and denoising by orientation-dependent tensor decomposition. Geophysics **2021**, 86, V107–V117. 10.1190/geo-2020-0070.1.

Cavalcante, Q.; Porsani, M.J. Low-rank seismic data reconstruction and denoising by CUR matrix decompositions. Geophysical Prospecting **2022**, 70, 362–376. 10.1111/1365-2478.13163.

Donoho, D.L. Compressed sensing. IEEE Transactions on Information Theory **2006**, 52, 1289–1306.

Duijndam, A.J.W.; Schonewille, M.A.; Hindriks, C.O.H. Reconstruction of band-limited signals, irregularly sampled along one spatial direction. Geophysics **1999**, 64, 524–538. 10.1190/1.1444559.

Eckart, C.; Young, G. The approximation of one matrix by another of lower rank. Psychometrika **1936**, 1, 211–218. 10.1007/BF02288367.

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. 10.1190/geo2014-0467.1.

Fomel, S. Seismic reflection data interpolation with differential offset and shot continuation. Geophysics **2003**, 68, 1–12. 10.1190/1.1567243.

Ford, W. Numerical linear algebra with applications: using MATLAB, 1 ed.; Academic Press: London, 2014.

Freire, S.L.M.; Ulrych, T.J. Application of singular value decomposition to vertical seismic profiling.

Geophysics **1988**, 53, 778–785. 10.1190/1.1442513.

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. 10.1190/geo2012-0038.1.

Gao, J.; Stanton, A.; Sacchi, M.D. Parallel matrix factorization algorithm and its application to 5D seismic reconstruction and denoising. Geophysics **2015**, 80, V173–V187. 10.1190/geo2014-0594.1.

Golub, G.H.; Van Loan, C.F. Matrix computations, 3 ed.; Johns Hopkins University Press: London, 1996.

Herrmann, F.J. Randomized sampling and sparsity: Getting more information from fewer samples. Geophysics **2010**, 75, WB173–WB187.

Huang, W.; Feng, D.; Chen, Y. De-aliased and de-noise Cadzow filtering for seismic data reconstruction. Geophysical Prospecting **2020**, 68, 553–571. 10.1111/1365-2478.12867.

Jia, Y.; Ma, J. What can machine learning do for seismic data processing? An interpolation application. Geophysics **2017**, 82, V163–V177. 10.1190/geo2016-0300.1.

Kaplan, S.T.; Naghizadeh, M.; Sacchi, M.D. Data reconstruction with shot-profile least-squares migration. Geophysics **2010**, 75, WB121–WB136. 10.1190/1.3478375.

Kaur, H.; Pham, N.; Fomel, S. Seismic data interpolation using deep learning with generative adversarial networks. Geophysical Prospecting **2021**, 69, 307–326. 10.1111/1365-2478.13055.

Kirlin, R.L. Vector subspaces. In Covariance analysis for seismic signal processing, 1 ed.; Kirlin, R.L.; Done, W.J., Eds.; Society of Exploration Geophysicists: Tulsa, 1999; chapter 4.

Kolda, T.G.; Bader, B.W. Tensor decompositions and applications. SIAM Review **2009**, 51, 455–500. 10.1137/07070111x.

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. 10.1190/geo2011-0399.1.

Kreimer, N.; Stanton, A.; Sacchi, M.D. Tensor completion based on nuclear norm minimization for 5D seismic data reconstruction. Geophysics **2013**, 78, V273–V284. 10.1190/geo2013-0022.1.

Liu, Y.; Fomel, S. Seismic data interpolation beyond aliasing using regularized nonstationary autoregression. Geophysics **2011**, 76, V69–V77. 10.1190/geo2010-0231.1.

Liu, B.; Sacchi, M.D. Minimum weighted norm interpolation of seismic records. Geophysics **2004**, 69, 1560–1568. 10.1190/1.1836829.

Menke, W. Geophysical data analysis: Discrete inverse theory, 3 ed.; Academic Press: Oxford, 2012.

Naghizadeh, M.; Sacchi, M.D. Seismic data reconstruction using multidimensional prediction filters. Geophysical Prospecting **2010a**, 58, 157–173. 10.1111/j.1365-2478.2009.00805.x.

Naghizadeh, M.; Sacchi, M.D. Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data. Geophysics **2010b**, 75, WB189–WB202. 10.1190/1.3509468.

Naghizadeh, M.; Sacchi, M.D. On sampling functions and Fourier reconstruction methods. Geophysics **2010c**, 75, WB137–WB151.

Naghizadeh, M.; Sacchi, M.D. Multidimensional de-aliased Cadzow reconstruction of seismic records. Geophysics **2013**, 78, A1–A5. 10.1190/geo2012-0200.1.

Naghizadeh, M. Seismic data interpolation and denoising in the frequency-wavenumber domain. Geophysics **2012**, 77, V71–V80. 10.1190/geo2011-0172.1.

Oboué, Y.A.S.I.; Chen, W.; Wang, H.; Chen, Y. Robust damped rank-reduction method for simultaneous denoising and reconstruction of 5D seismic data. Geophysics **2021**, 86, V71–V89. 10.1190/geo2020-0032.1.

Oropeza, V.; Sacchi, M. Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis. Geophysics **2011**, 76, V25–V32. 10.1190/1.3552706.

Porsani, M.J. Seismic trace interpolation using half-step prediction filters. Geophysics **1999**, 64, 1461–1467. 10.1190/1.1444650.

Qin, F.; Leger, P.C.; Ren, J.; Aleksic, V.; Rowe, R.W.; Zainaldin, S.; Hadab, S.A. A robust implementation and application of antileakage Fourier transform interpolation. The Leading Edge **2018**, 37, 538–543. 10.1190/tle37070538.1.

Spitz, S. Seismic trace interpolation in the F-X domain. Geophysics **1991**, 56, 785–794. 10.1190/1.1443096.

Stolt, R.H. Seismic data mapping and reconstruction. Geophysics **2002**, 67, 890–908. 10.1190/1.1484532.

Strang, G. Linear Algebra and Its Applications, 4 ed.; Cengage Learning: Andover, 2018.

Tomasi, G.; Bro, R. PARAFAC and missing values. Chemometrics and Intelligent Laboratory Systems **2005**, 75, 163–180. 10.1016/j.chemolab.2004.07.003.

Trad, D. Interpolation and multiple attenuation with migration operators. Geophysics **2003**, 68, 2043–2054. 10.1190/1.1635058.

Trad, D. Five-dimensional interpolation: Recovering from acquisition constraints. Geophysics **2009**, 74, V123–V132. 10.1190/1.3245216.

Trickett, S.; Burroughs, L. Prestack rank-reduction-based noise suppression. CSEG **2009**, 34, 25–31.

Trickett, S.; Burroughs, L.; Milton, A.; Walton, L.; Dack, R. Rank-reduction-based trace interpolation. 80th Annual Int. Meeting, SEG, Expanded Abstracts **2010**, pp. 3829–3833. 10.1190/1.3513645.

Ulrych, T.J.; Sacchi, M.D.; Freire, S.L.M. Eigenimage processing of seismic sections. In Covariance analysis for seismic signal processing, 1 ed.; Kirlin, R.L.; Done, W.J., Eds.; Society of Exploration Geophysicists: Tulsa, 1999; chapter 12.

Vermeer, G.J.O. 3D seismic survey design, 2 ed.; Society of Exploration Geophysicists: Tulsa, 2012. 10.1190/1.9781560803041.

Wang, B.; Zhang, N.; Lu, W.; Wang, J. Deep-learning-based seismic data interpolation: A preliminary result. Geophysics **2019**, 84, V11–V20. 10.1190/geo2017-0495.1.

Xu, S.; Zhang, Y.; Pham, D.; Lambaré, G. Antileakage Fourier transform for seismic data regularization. Geophysics **2005**, 70, V87–V95. 10.1190/1.1993713.

Zhang, H.; Zhang, H.; Zhang, J.; Hao, Y.; Wang, B. An anti-aliasing POCS interpolation method for regularly undersampled seismic data using curvelet transform. Journal of Applied Geophysics **2020**, 172, 1–11. 10.1016/j.jappgeo.2019.103894.

Zwartjes, P.M.; Sacchi, M.D. Fourier reconstruction of nonuniformly sampled, aliased seismic data. Geophysics **2007**, 72, V21–V32. 10.1190/1.2399442.
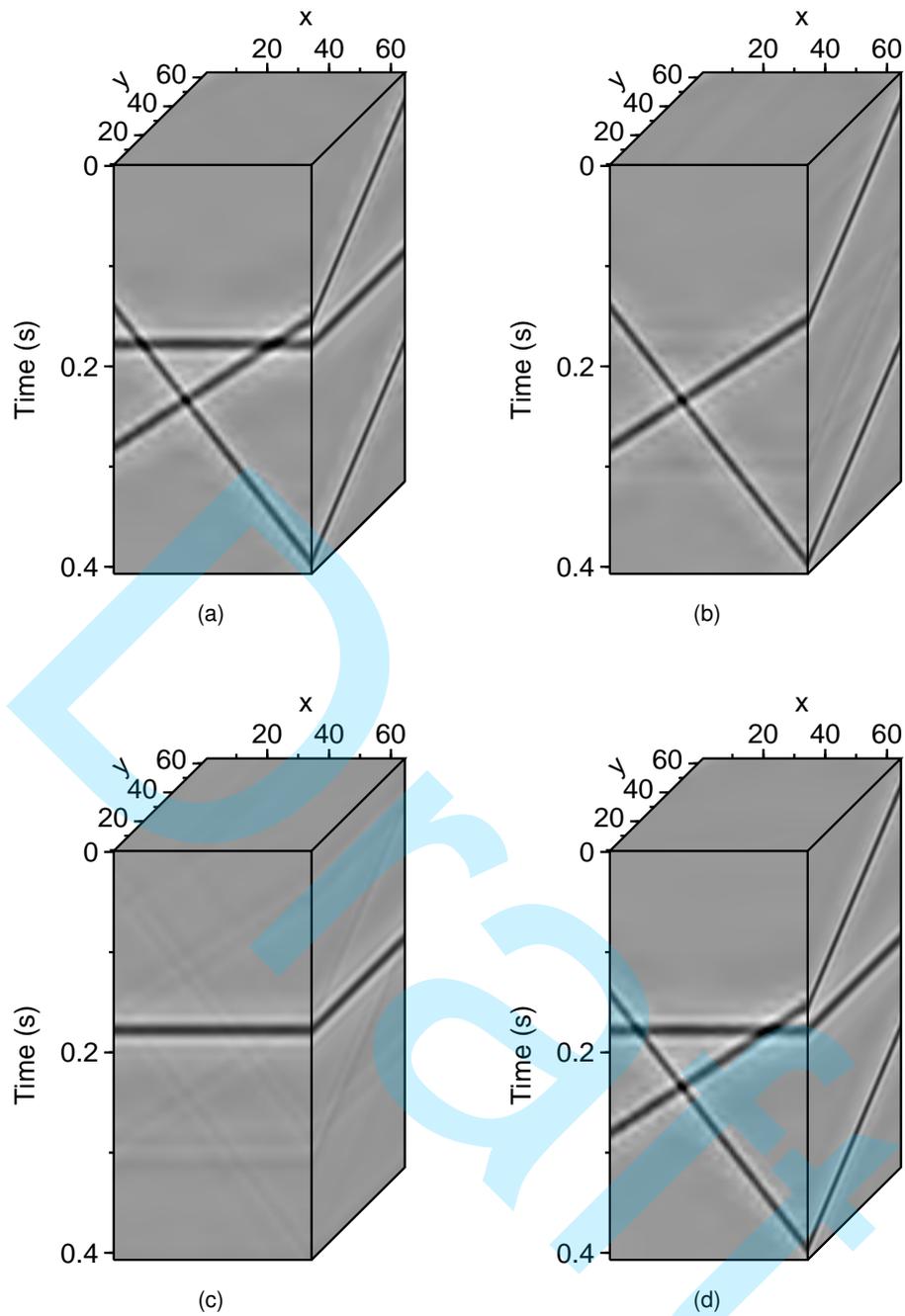
Figure 1. 3D synthetic data with linear events. (a) Original volume. (b) First weighted eigenvolume. (c) Second weighted eigenvolume. (d) Low-rank approximation using the first two eigenvolumes (b) + (c).
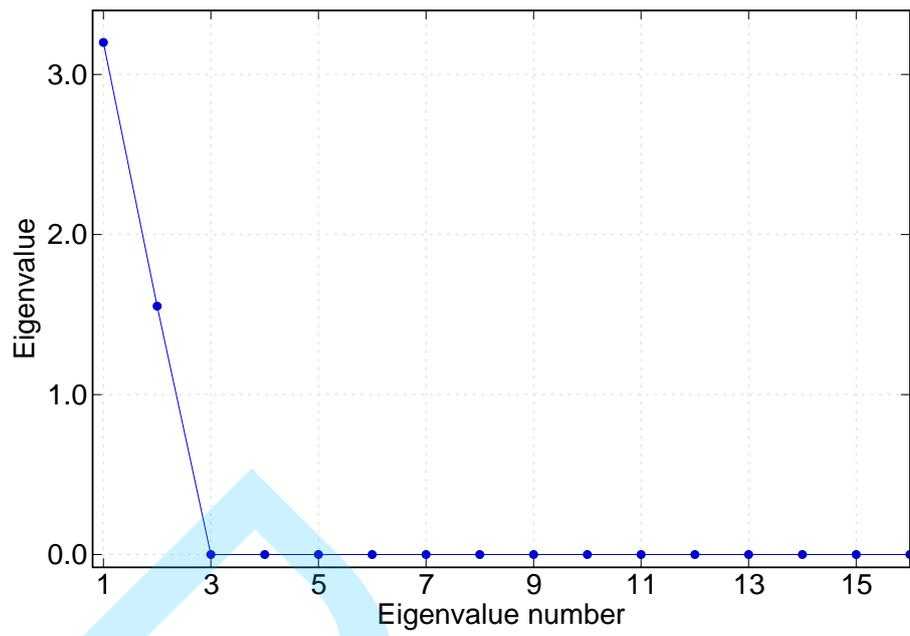
Figure 2. Eigenvalues corresponding to the 3D synthetic volume with linear events for frequency $f = 29.3\,$Hz.
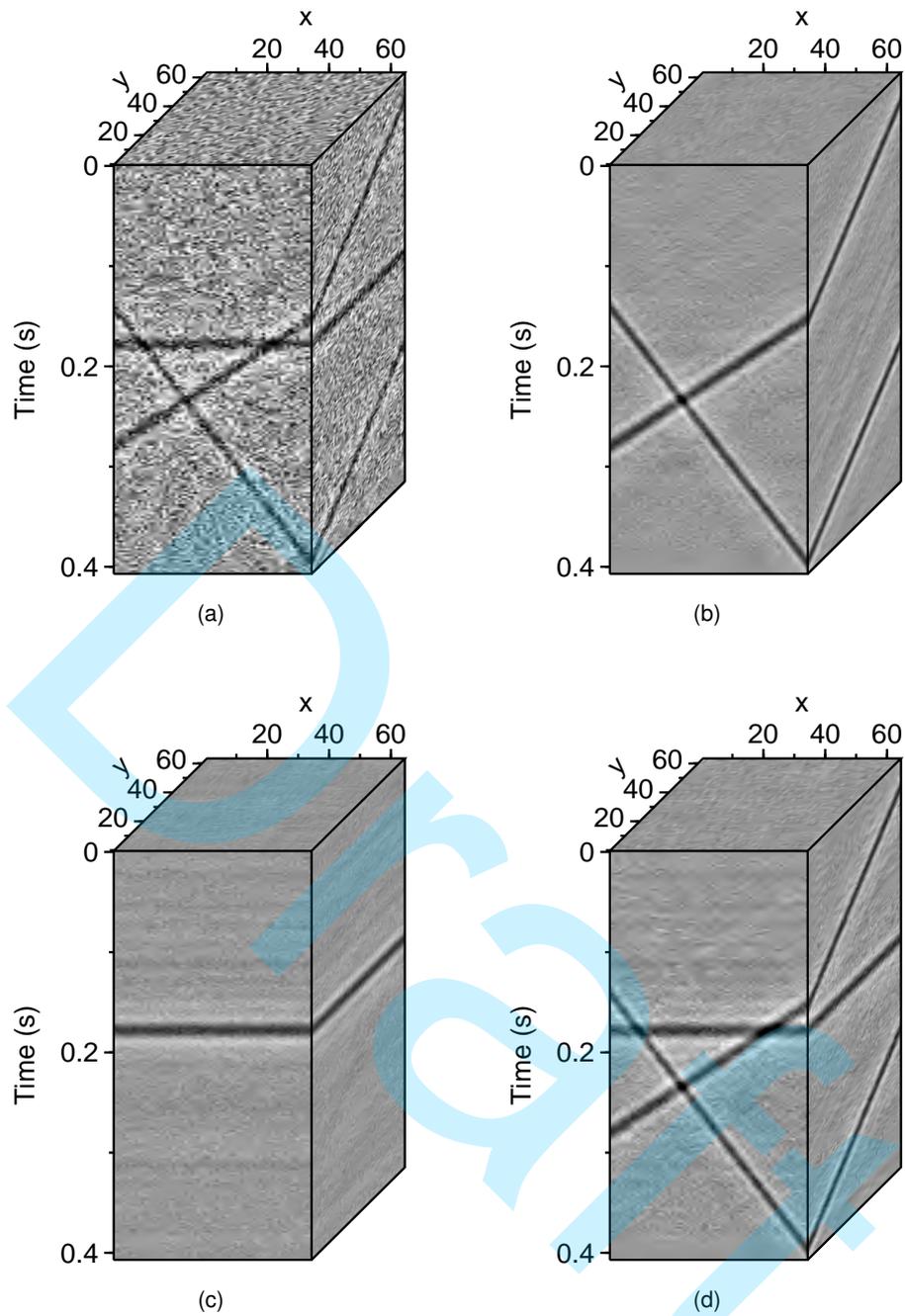
Figure 3. 3D synthetic data with linear events and additive Gaussian noise. (a) Original volume. (b) First weighted eigenvolume. (c) Second weighted eigenvolume. (d) Low-rank approximation using the first two eigenvolumes (b) + (c).
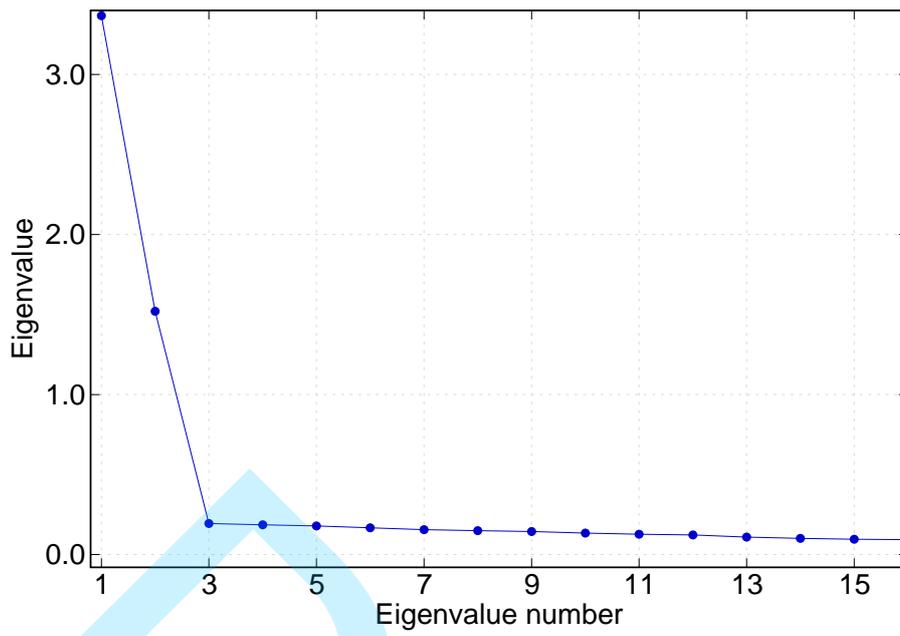
Figure 4. Eigenvalues corresponding to the 3D synthetic volume with linear events and additive Gaussian noise, for frequency $f = 29.3$ Hz.
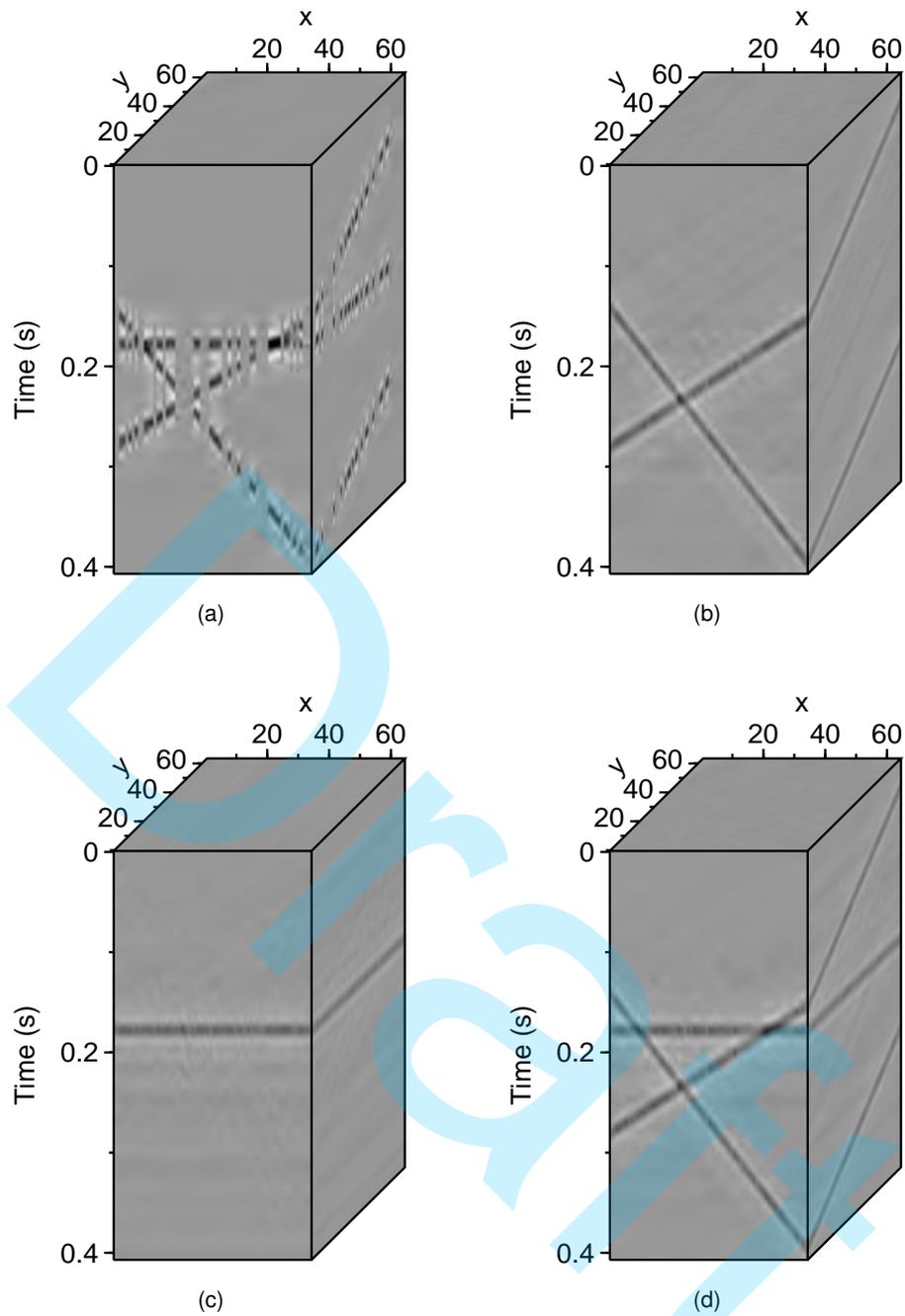
Figure 5. 3D synthetic data with linear events. (a) Original volume with nearly 50% of missing traces. (b) First weighted eigenvolume. (c) Second weighted eigenvolume. (d) Low-rank approximation using the first two eigenvolumes (b) + (c).
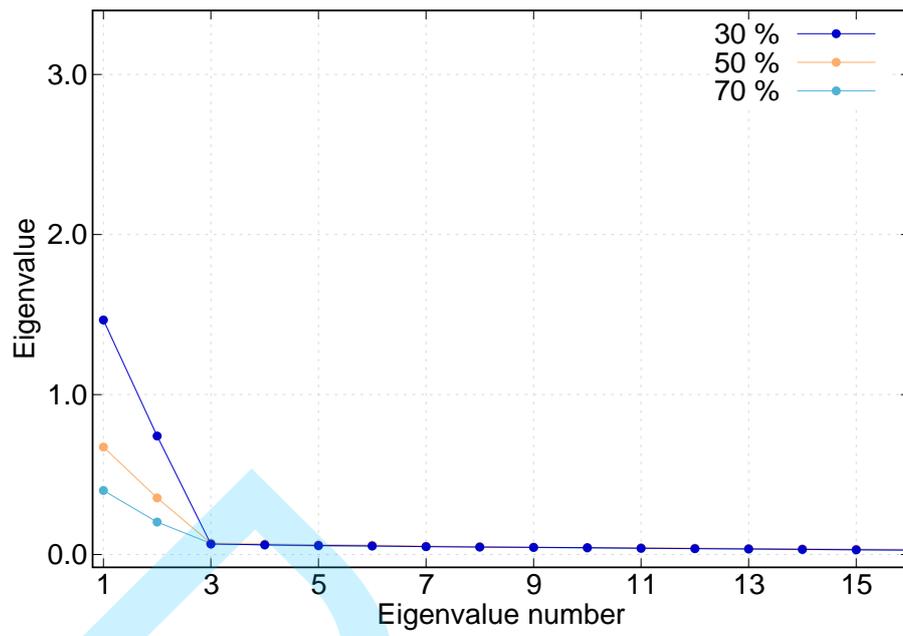
Figure 6. Eigenvalues corresponding to the 3D synthetic volume with linear events and three decimation percentages, for frequency $f = 29.3$ Hz.
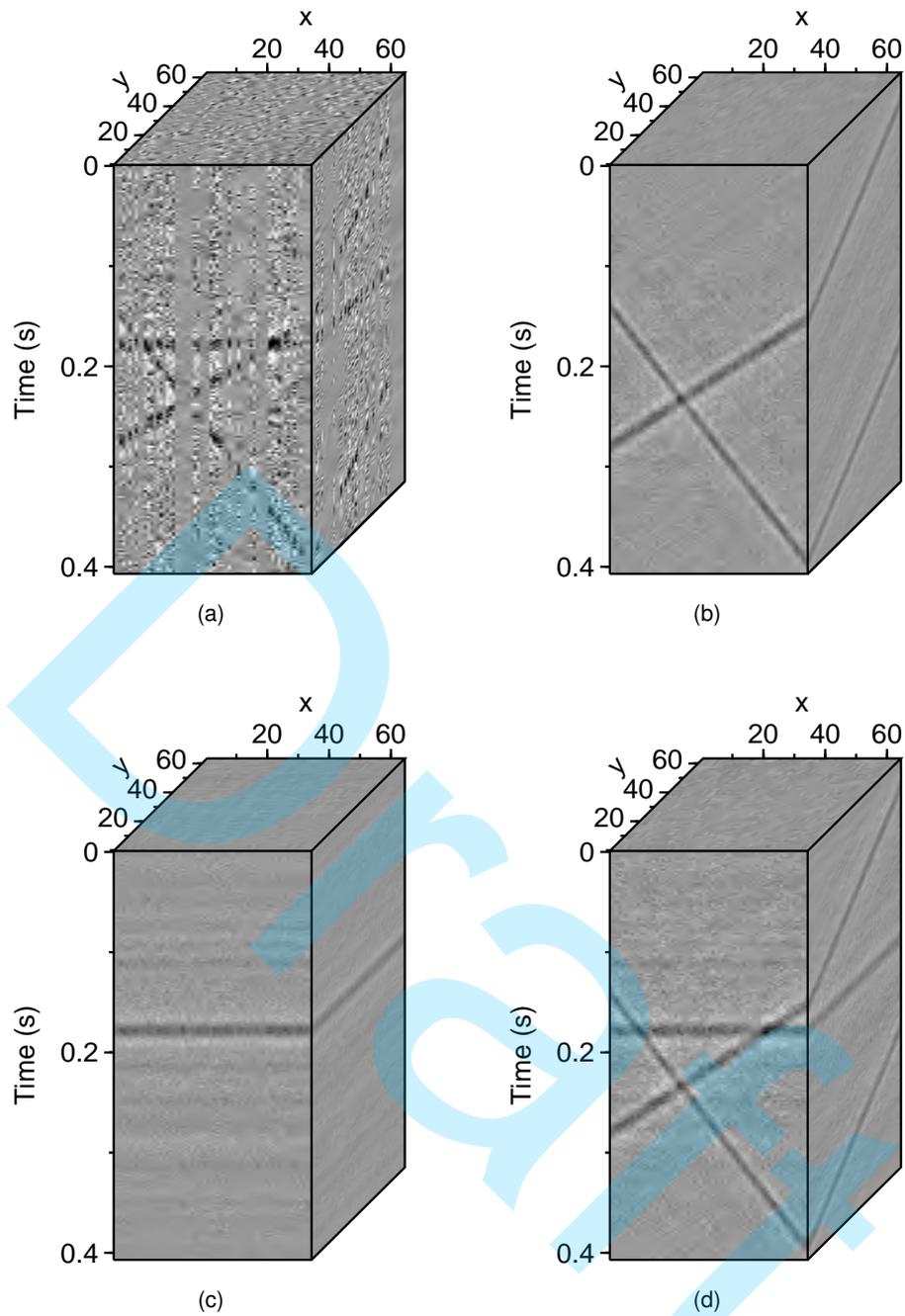
Figure 7. 3D synthetic data with linear events and additive Gaussian noise. Nearly 50% of the traces were removed in (a). (b) First weighted eigenvolume. (c) Second weighted eigenvolume. (d) Low-rank approximation using the first two eigenvolumes (b) + (c).
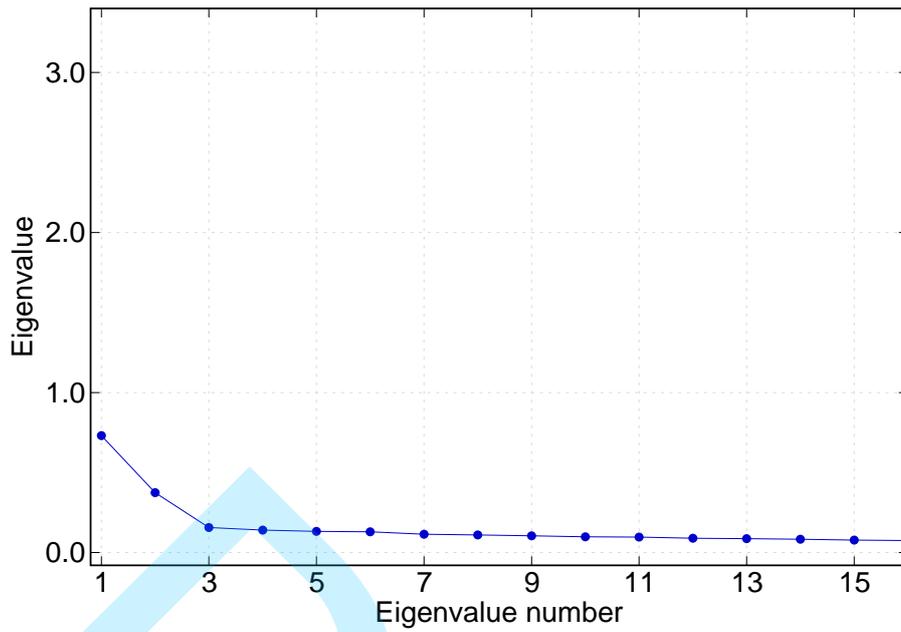
Figure 8. Eigenvalues corresponding to the irregular 3D synthetic volume with linear events and additive noise, for frequency $f = 29.3$ Hz.
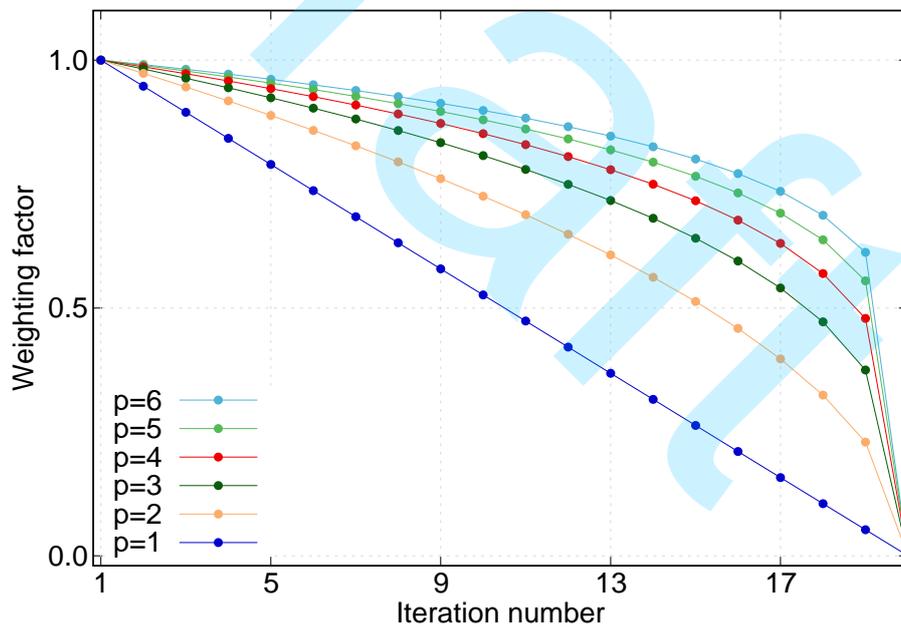
Figure 9. Weighting factor a as a function of the iteration number for different values of p (equation 19).
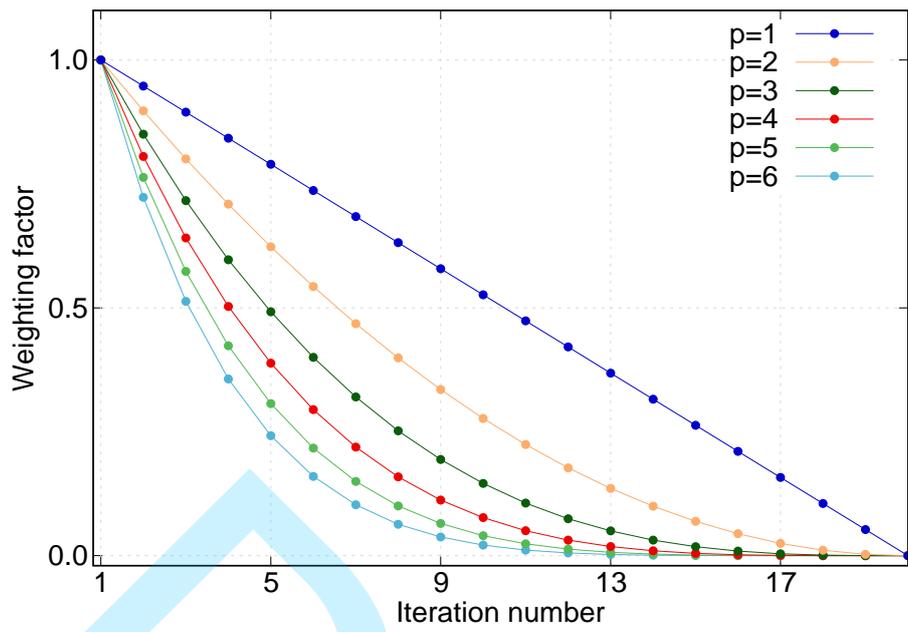
Figure 10. Weighting factor a as a function of the iteration number for different values of p (equation 20).
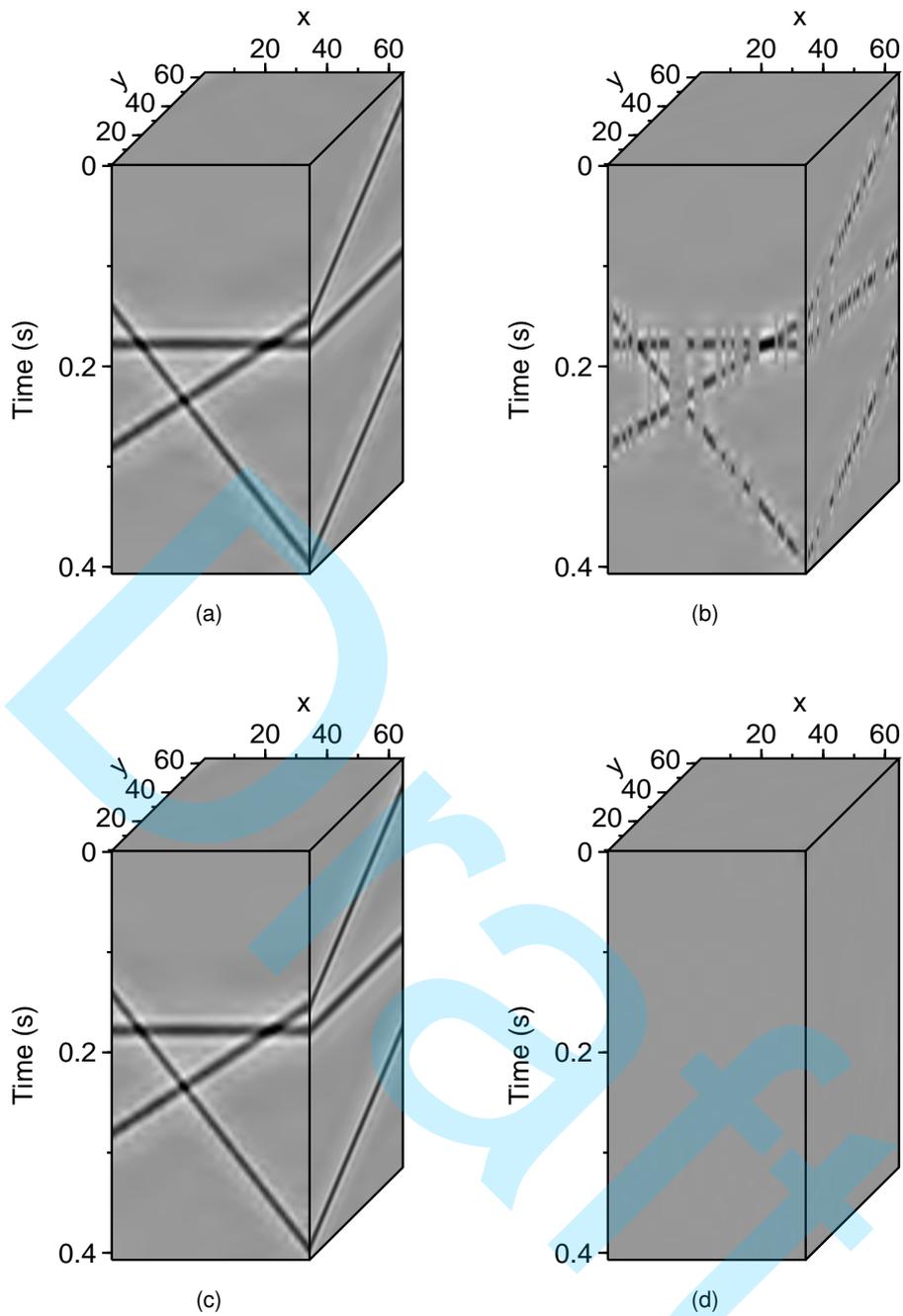
Figure 11. Reconstruction of the 3D synthetic data with linear events. (a) Original volume. (b) Decimated volume. (c) Recovery result using $r = 2$. $Q = 46.9$ dB. (d) Reconstruction error (c) - (a).
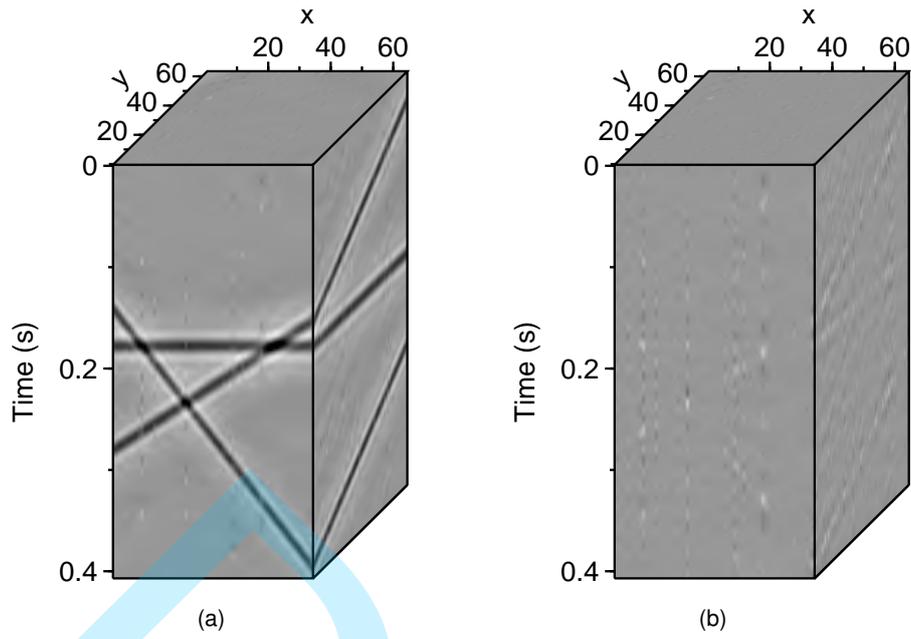
Figure 12. Reconstruction of the 3D synthetic data with linear events. (a) Recovery result using $r = 5$. $Q = 19.2$ dB. (d) Reconstruction error (a) - 11(a).
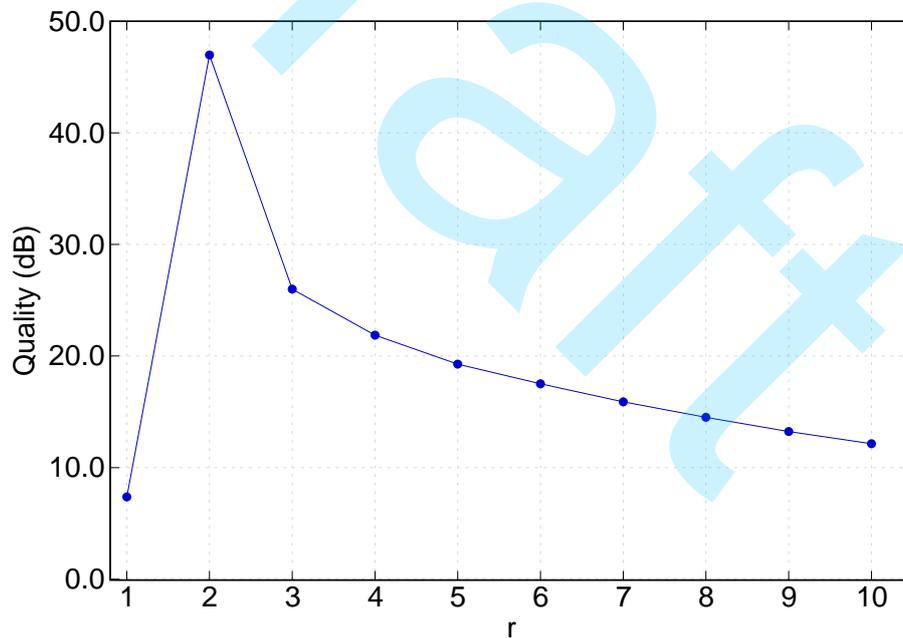
Figure 13. Reconstruction quality as a function of the rank r for the 3D synthetic volume with linear events.
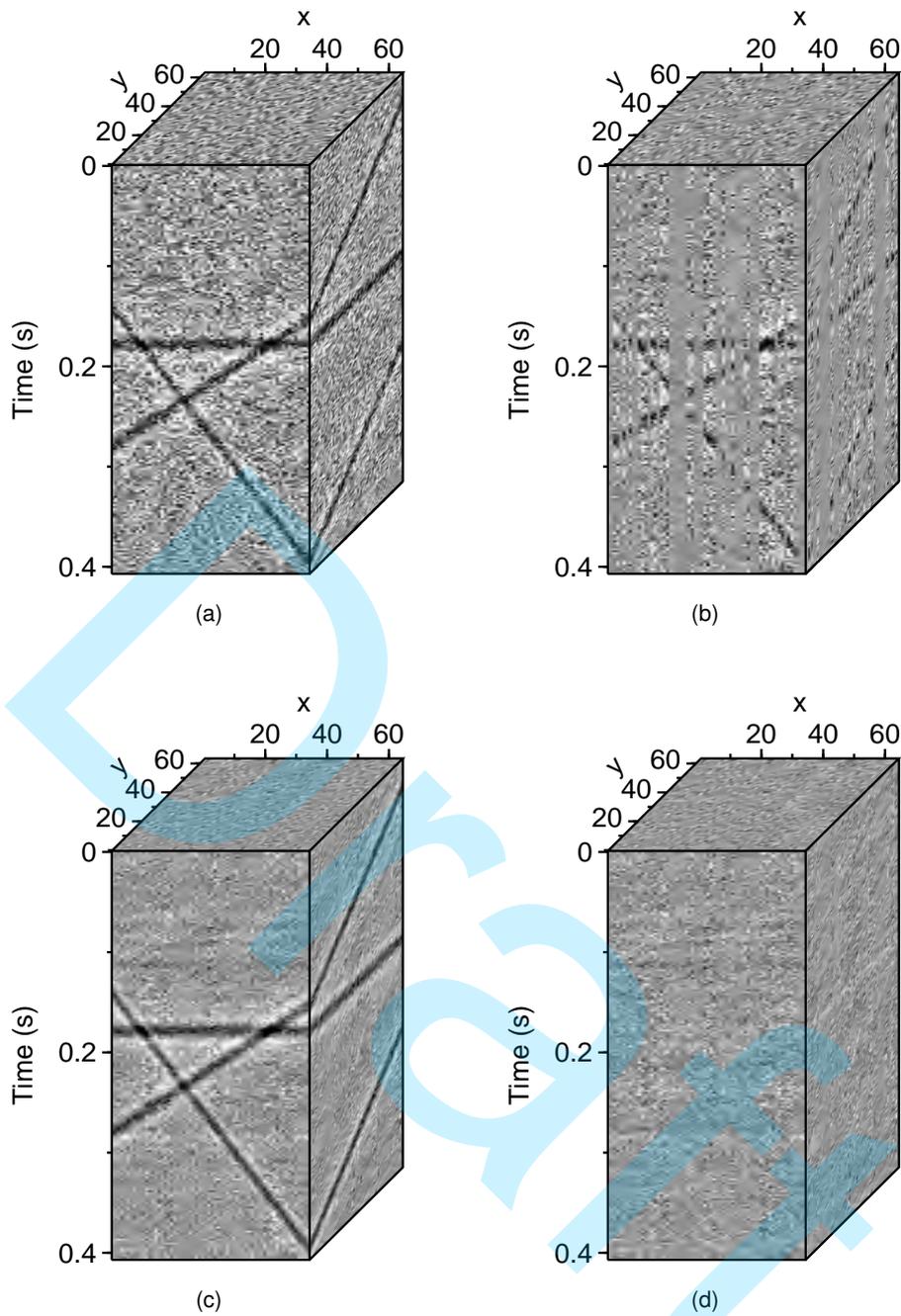
Figure 14. Reconstruction of the 3D synthetic data with linear events and additive Gaussian noise. (a) Original noisy volume. (b) Decimated volume. (c) Recovery result using $r = 2$ and $a = 0.5$. $Q = -0.3$ dB. (d) Reconstruction error (c) - 11(a).
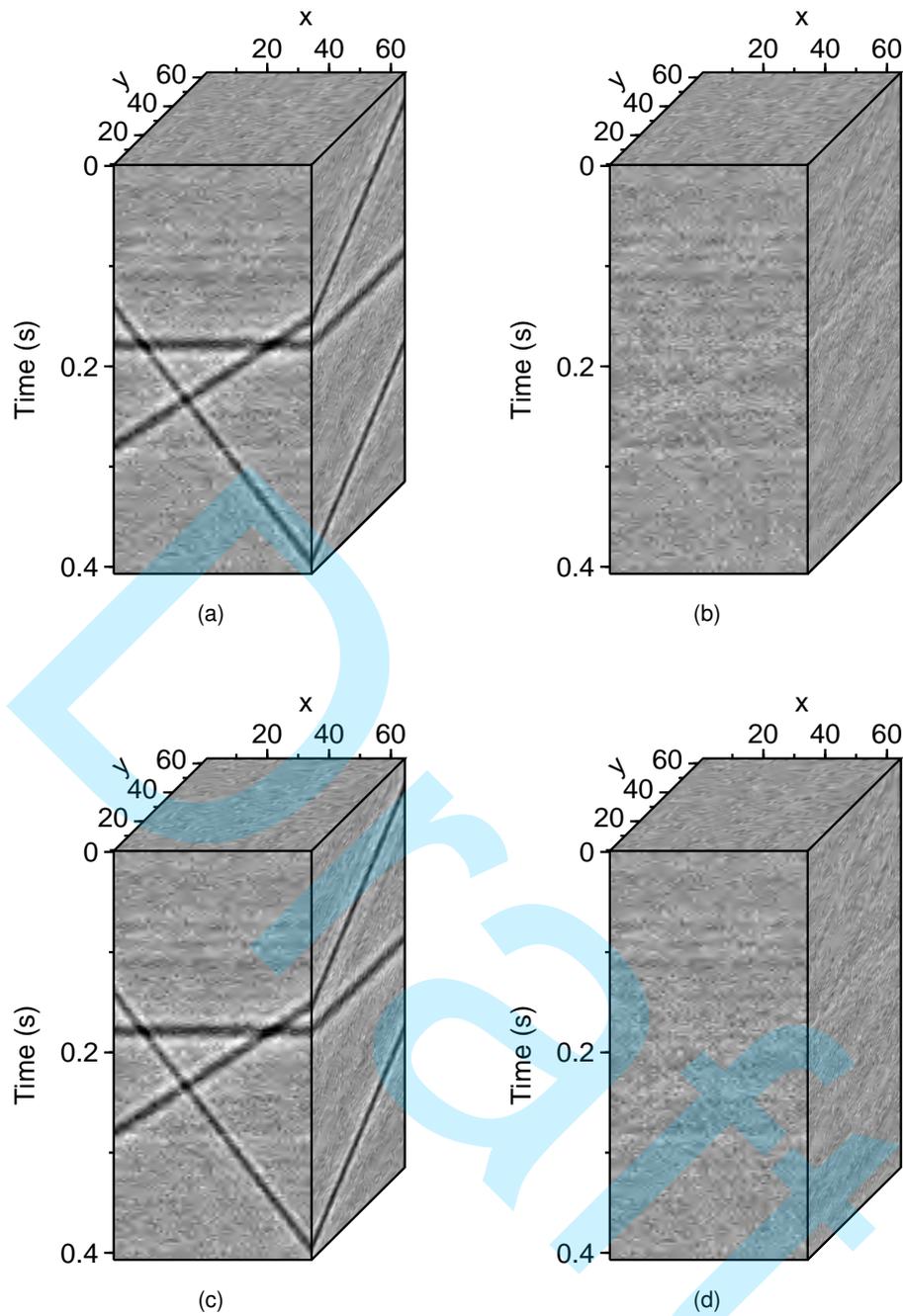
Figure 15. Reconstruction of the 3D synthetic data with linear events and additive Gaussian noise. (a) Recovery result using $r = 2$ and $p = 1$ in equation 19 (fractional exponent). $Q = 2.7\,$dB. (b) Reconstruction error (a) - 10(a). (c) Recovery result using $r = 2$ and $p = 3$. $Q = 2.4\,$dB. (d) Reconstruction error (c) - 11(a).
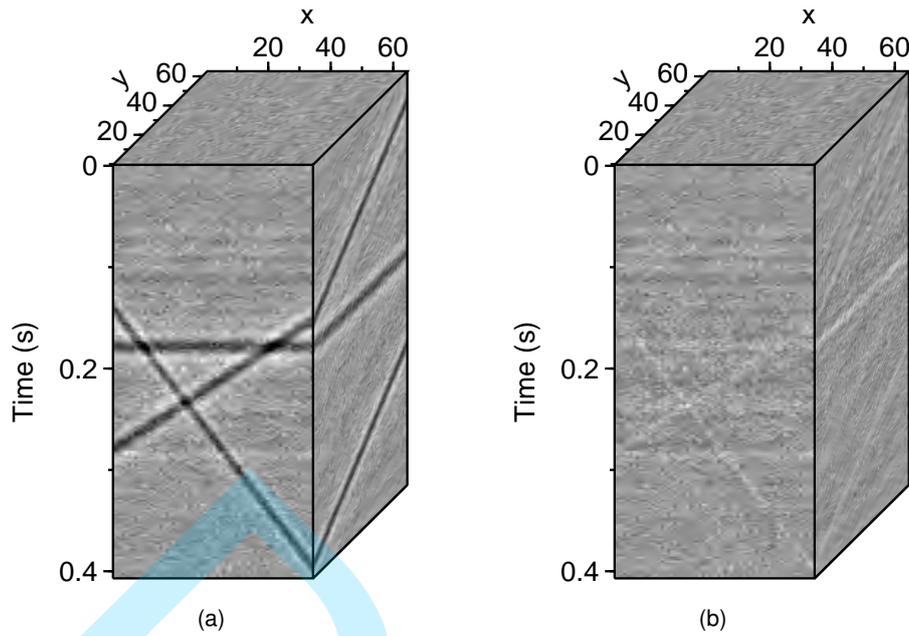
Figure 16. Reconstruction of the 3D synthetic data with linear events and additive Gaussian noise. (a) Recovery result using $r = 2$ and $p = 3$ in equation 20 (integer exponent). $Q = 3.3$ dB. (b) Reconstruction error (a) - 11(a).
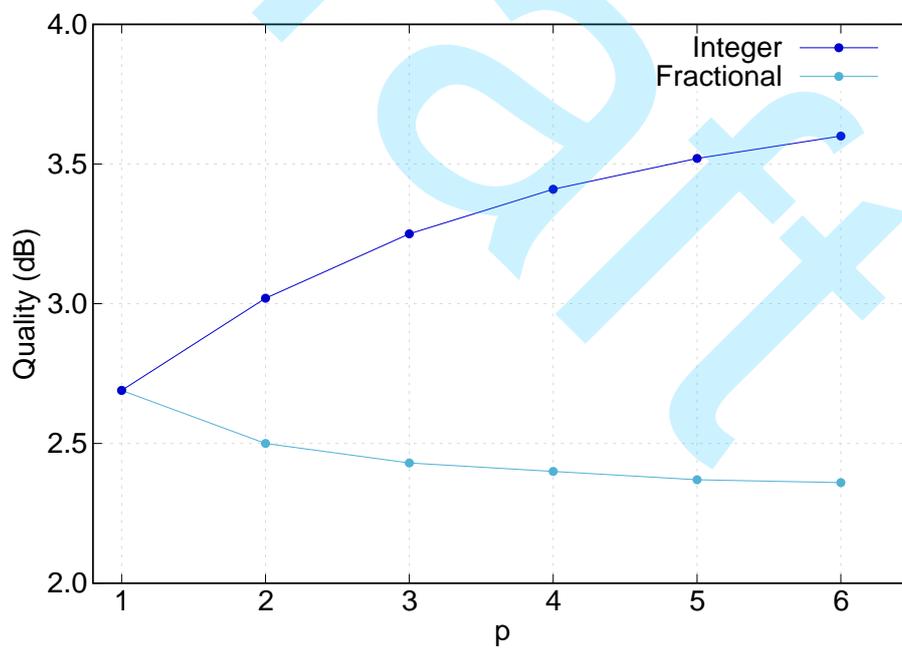
Figure 17. Reconstruction quality as a function of the parameter p in equations 19 (fractional exponent) and 20 (integer exponent), for the 3D synthetic volume with linear events and additive Gaussian noise.
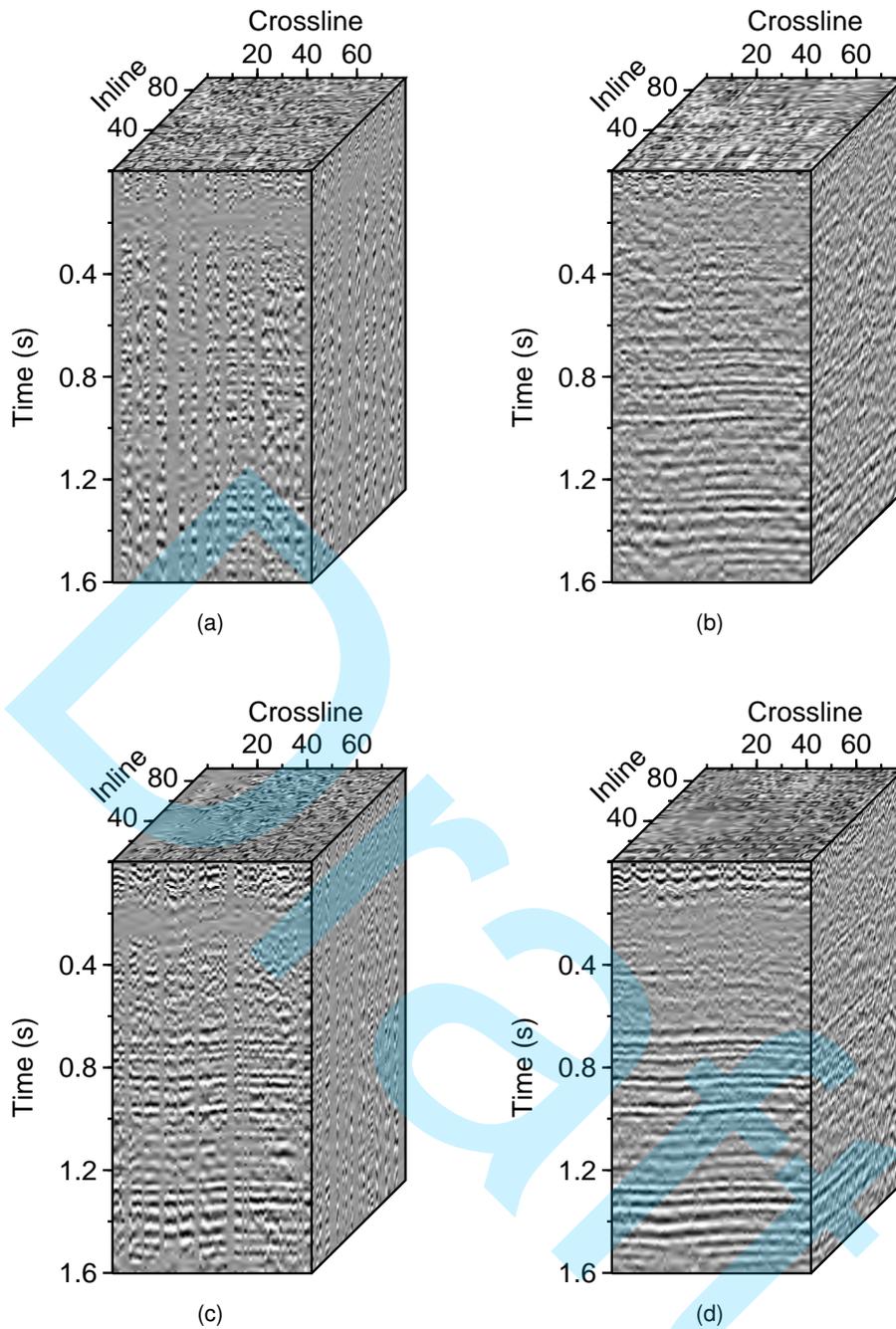
Figure 18. Reconstruction of 3D common-offset volumes using $r = 3$ in spatial windows of size 21×21 traces. (a) Irregular volume corresponding to offset bin number 3, centered at 300 m. (b) Recovery result. (c) Irregular volume corresponding to offset bin number 4, centered at 420 m. (d) Recovery result.
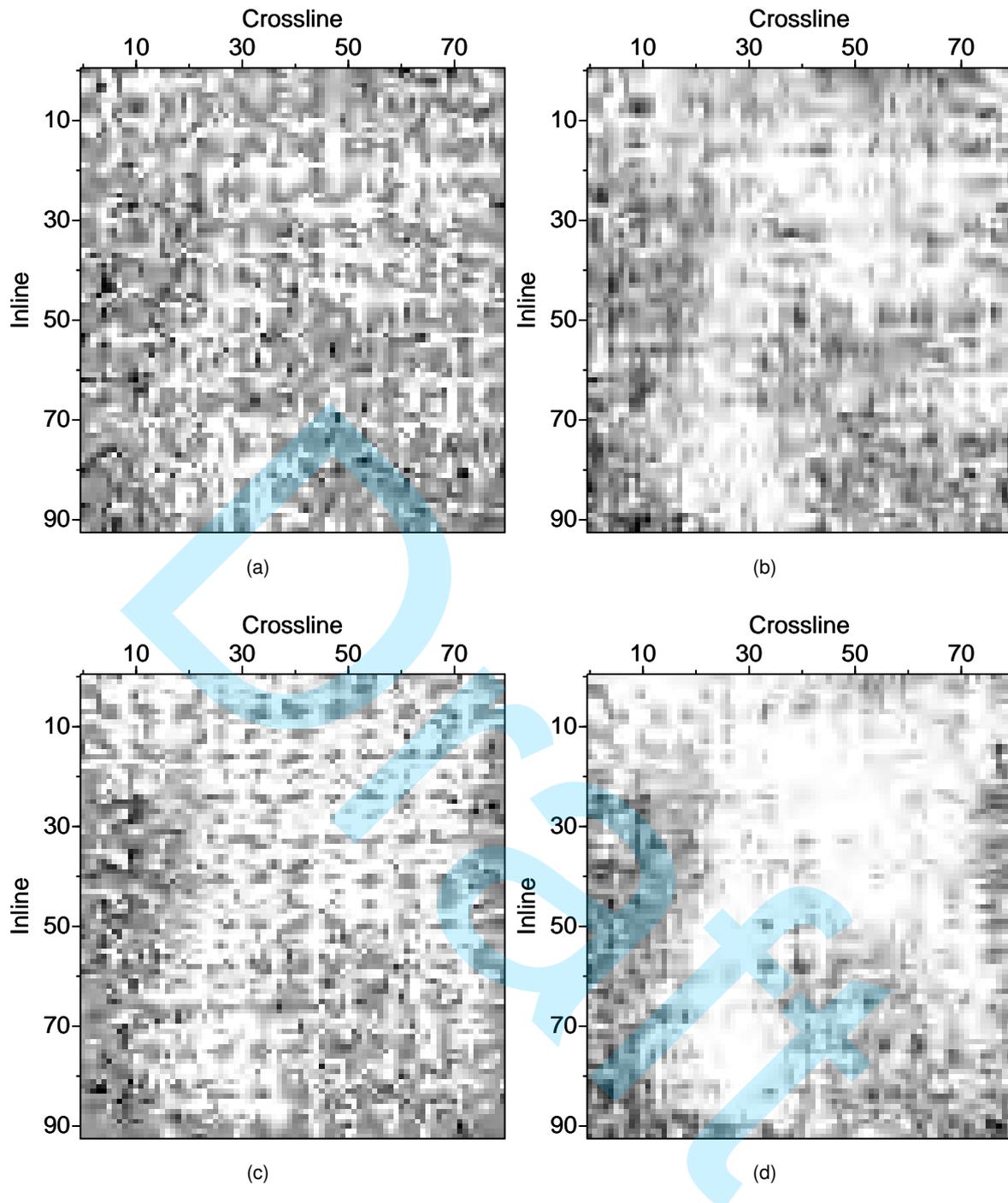
Figure 19. Time slices at 1352 ms of common-offset volumes. The grayscale is the same as Figure 18. (a) Original time slice of offset bin number 3, centered at 300 m. (b) Recovered time slice. (c) Original time slice of offset bin number 4, centered at 420 m. (d) Recovered time slice.