GEOPHYSICS, VOL. 80, NO. 4 (JULY-AUGUST 2015); P. V83–V95, 11 FIGS.
10.1190/GEO2014-0467.1

5D seismic data completion and denoising using a novel class of tensor decompositions

Gregory Ely[1], Shuchin Aeron[2], Ning Hao[3], and Misha E. Kilmer[3]

ABSTRACT

We have developed a novel strategy for simultaneous interpolation and denoising of prestack seismic data. Most seismic surveys fail to cover all possible source-receiver combinations, leading to missing data especially in the midpoint-offset domain. This undersampling can complicate certain data processing steps such as amplitude-variation-with-offset analysis and migration. Data interpolation can mitigate the impact of missing traces. We considered the prestack data as a 5D multidimensional array or otherwise referred to as a *5D tensor*. Using synthetic data sets, we first found that prestack data can be well approximated by a low-rank tensor under a recently proposed framework for tensor singular value decomposition (tSVD). Under this low-rank assumption, we proposed a complexity-penalized algorithm for the recovery of missing traces and data denoising. In this algorithm, the complexity regularization was controlled by tuning a single regularization parameter using a statistical test. We tested the performance of the proposed algorithm on synthetic and real data to show that missing data can be reliably recovered under heavy downsampling. In addition, we demonstrated that compressibility, i.e., approximation of the data by a low-rank tensor, of seismic data under tSVD depended on the velocity model complexity and shot and receiver spacing. We further found that compressibility correlated with the recovery of missing data because high compressibility implied good recovery and vice versa.

INTRODUCTION

An ideal seismic survey would cover all possible source and receiver combinations. These geometries rarely occur due to financial and physical constraints. In particular, the shot and receiver positions are planned prior to acquisition but the true sampling locations usually deviate from the originally intended ones resulting in less desirable midpoint and azimuth sampling. Moreover, the midpoint-offset-azimuth domain is also not fully sampled due to the intrinsic nature of seismic acquisition geometries. These effects are more pronounced in land surveys. For methods that rely on gridded data, the extent of undersampling depends on the manner in which the data are binned. Also note that even under complete acquisition for a given geometry, the seismic wavefield may still be undersampled in the domain in which other processes are carried out. This spatial undersampling adversely affects later processing steps such as migration and quantitative interpretation (QI) studies such as amplitude versus angle analysis (Sacchi and Liu, 2005; Hunt et al., 2010). Recovery of prestack traces can mitigate the impact of incomplete sampling. In this paper, we examine methods based in linear algebra for recovery of missing traces.

Prestack seismic data can be viewed as a fifth-order tensor consisting of one time or frequency dimension and four spatial dimensions describing the location of the sources and the receivers on the surface, either in the original source-receiver (r_x, r_y, s_x, s_y) coordinate frame or in terms of midpoints and offsets (x, y, h_x, h_y). Some of the current approaches, such as in Trad (2009) and Kreimer and Sacchi (2012b), work the *f-x* domain in which the data completion is carried out separately over the 4D slices corresponding to each frequency.

Prestack seismic reconstruction techniques typically exploit the fact that under some assumptions, the fully sampled seismic wavefield is highly redundant and can be compressed under an appropriate representation. Most of these methods fall into two categories: techniques that explicitly define a transform domain and methods that do not define a basis, instead representing the seismic data as a low-rank matrix or tensor. Numerous methods exploit compactness in the frequency-space (*f-x*) domain (Duijndam et al., 1999; Hindriks and Duijndam, 2000; Liu and Sacchi, 2004; Xu et al., 2005; Zwartjes and Sacchi, 2007; Trad, 2009; Curry,

Manuscript received by the Editor 1 October 2014; revised manuscript received 26 March 2015; published online 8 June 2015.
[1]Massachusetts Institute of Technology, Earth Resources Laboratory, Cambridge, Massachusetts, USA. E-mail: elyg@mit.edu.
[2]Tufts University, Department of Electrical Engineering, Medford, Massachusetts, USA. E-mail: shuchin@ece.tufts.edu.
[3]Tufts University, Department of Mathematics, Medford, Massachusetts, USA. E-mail: ning.hao@tufts.edu; misha.kilmer@tufts.edu.

2010) and other domains such as the curvelet (Hennenfent and Herrmann, 2006; Herrmann and Hennenfent, 2008; Naghizadeh and Sacchi, 2010a), seislets (Fomel and Liu, 2010), and Radon (Kabir and Verschuur, 1995). Other methods exploit the statistical compactness or the predictably of a signal in the *f-x* domain (Spitz, 1991; Gulunay, 2003; Naghizadeh and Sacchi, 2007, 2010b). In addition to methods that either impose sparsity explicitly or compactness through thresholding in a basis, low-rank matrix methods do not define a single basis and instead impose compactness on the signal's subspace. In these low-rank methods, the 5D seismic data are embedded into a 2D Hankel matrix in which the true noiseless and complete signal is assumed to be of low rank (Trickett et al., 2010). Iterative thresholding and imposing sparsity on the singular values of the matrix enforces the reduced rank of the data, requiring numerous costly singular value decompositions (SVDs). Other variations enforce the low rank through less expensive matrix decomposition, completely avoiding the calculation of the SVD (Gao et al., 2011, 2013; Kumar et al., 2013) or applying the SVD to a randomized subset of the original matrix (Oropeza and Sacchi, 2010, 2011). In contrast to the aforementioned methods, other methods use a priori velocity information and the wave equation to interpolate missing traces (Ronen, 1987; Stolt, 2002; Fomel, 2003; Kaplan et al., 2010), thereby incorporating actual physical information and principles governing the system. See Stanton and Sacchi (2013) for a detailed survey of current reconstruction methods.

More recently, methods for seismic completion have exploited multilinear algebraic approaches based on tensor rank minimization. Unlike matrices, for which the rank is uniquely defined, for third- or higher order tensors, the rank heavily depends on the choice of tensor factorization and rank penalization can encourage different structures of the tensor. SVD-like tensor decompositions can be divided into three classes: canonical decomposition (CANDECOMP, CP) (Kolda and Bader, 2009), Tucker or higher-order SVD (HOSVD) (Lathauwer et al., 2000), and hierarchical tucker (HT) (Grasedyck, 2010). All of these decompositions have been applied to the problem of 5D seismic data reconstruction with good performance on real and synthetic data. Seismic data completion using tensor rank minimization under HOSVD is proposed by Kreimer and Sacchi (2012b) and Kreimer et al. (2013). Although the performance was shown to be good, the proposed algorithm requires preselection of the truncation ranks along each dimension for effective interpolation. No systematic approach has been proposed to select these truncation ranks. One way to select these truncation ranks is to perform tests on a small piece of the data set and then use these parameters for the rest of the data. Although this approach is data adaptive, it is not clear if the ranks should remain constant over the entire data set. If this is not the case, then it may require ground truth and model selection for each section of the data. Kreimer et al. (2013) propose seismic data completion based on nuclear norm minimization. This approach flattens the data into various modes, which may be suboptimal because it destroys the multidimensional structure of the data. Recently, Zhang et al. (2014) demonstrate this information loss for the problem of 3D and 4D video data completion from missing pixels. Even though video data and seismic data are very different data types, we expect that retaining the multidimensional structure of the data should also help with recovery in the seismic case. Nevertheless, it is important to note that because we do not compare our method with that in Kreimer et al. (2013), we cannot claim that our approach is superior to theirs. Extensive comparison of our approach with those of Kreimer et al. (2013) is a subject of future work and will be undertaken in the future. In Silva and Herrmann (2013), the authors use the HT decomposition for data completion with a novel algorithm that exploits the manifold structure of tensors with fixed HT rank. Similar to the HOSVD case, the performance heavily depends on the prior knowledge of the dimension tree. Furthermore, it is computationally difficult to search for the best dimension tree for a given data set. In Trickett et al. (2013), CANDECOMP (CP)-type factorization is used for data completion. Although this approach requires selecting only a single parameter, namely, the CP rank, finding the CP decomposition is computationally challenging, and in some cases, the best r-rank approximations may not even exist (Silva and Lim, 2008).

We exploit a novel tensor decomposition, tensor singular value decomposition (tSVD), for completion of prestack seismic data. Kilmer and Martin (2011a) develop tSVD for third-order tensors and Martin et al. (2013) extend it to higher orders. Ely et al. (2013) present the initial results of this approach for seismic applications, and Zhang et al. (2014) present the accompanying results for computer vision applications. For 5D synthetic seismic data, we show that the prestack seismic data are compressible; i.e., they have low informational complexity expressed as the low tensor nuclear norm (TNN) in the tSVD domain. This compressibility stems from the convolutional structure of the t-product because the tSVD-based representation is particularly efficient at describing shift and scaling operations between traces. Hence, the data can be reliably recovered under limited sampling using the appropriate complexity penalized (CP) algorithm.

The tSVD differs from the HOSVD, CANDECOMP, and HT formats because the tSVD retains a tensor's orientation information while remaining computationally efficient. The TNN derived from tSVD differs significantly from the sum of nuclear norms approach used by Gandy et al. (2011) and Kreimer et al. (2013), where the summation is over the nuclear norms of all the different mode unfoldings of the tensor. Our approach computes the TNN, as derived from the tSVD, by first applying the Fourier transform recursively along the different dimensions and then summing the nuclear norms of the frontal slices of the resulting tensor. In contrast to the approach in Kreimer et al. (2013), the TNN of a tensor as computed from the tSVD depends significantly on the ordering of dimensions. This is also in contrast to the HOSVD approach used by Kreimer and Sacchi (2012a) and the approaches based on HT as used by Silva and Herrmann (2013), which are orientation independent. For example, two tensors containing the same data, one indexed as (t, r_x, r_y, s_x, s_y) and the other indexed as (r_x, r_y, s_x, s_y, t), would have two different TNNs as computed from tSVD. The approach based on the sum of nuclear norms of mode unfoldings is independent of orientation unless an a priori weighted sum is applied to each of the unfolded matrices. This weighting, when unknown or set to be uniform, can impose low-rank along foldings that should be of high rank or vice versa. Similar considerations also apply to the cases when using the HOSVD and HT decompositions, in which the choice of truncation ranks is orientation independent.

We organize the remainder of the article as follows: We first outline the mathematical problem formulation and review the background of the t-product and tSVD. We first perform a detailed case study of prestack seismic data compressibility under two extreme cases: a set of Born scatters and finite-difference data sets for

two different velocity models and two survey geometries. We then apply the proposed algorithm for a variety of sampling rates and plot the reconstruction performance. We then apply the simultaneous denoising and reconstruction algorithm to a real data example. Finally, we provide conclusions and future research directions.

METHOD

Notation: We will denote vectors by boldface small letters $\mathbf{x}$, matrices by boldface capital letters $\mathbf{X}$, and tensors (of order >2) by capital boldface script letters $\mathcal{X}$. Throughout the text, we use MATLAB notation and indexing to denote the elements and fibers of these vectors, matrices, and tensors. For example, $\mathcal{X}(:,:,1)$ will denote the first frontal slice of $\mathcal{X}$, $\mathcal{X}(i,j,:)$ will denote a tensor fiber into the page, and $\mathbf{X}(:,1)$ will denote the first column of $\mathbf{X}$ and so on.

Problem setup

The true seismic data $\mathcal{M}$ are spatially undersampled either in the source-receiver or in the midpoint-offset domain. The measurement $\mathcal{Y}$ can be represented by a linear operator $\mathcal{A}$, resulting in the sparse observed data $\mathcal{Y}$ under additive noise $\mathcal{N}$ as

$$\mathcal{Y} = \mathcal{A}(\mathcal{X}) + \mathcal{N}. \tag{1}$$

The seismic data completion problem becomes reliably estimating $\mathcal{X}$ from $\mathcal{Y}$ under the sampling operator $\mathcal{A}$. Because the number of observed measurements is significantly less than the number of elements in $\mathcal{X}$, the inverse problem is severely ill posed and cannot be solved directly without placing constraints on $\mathcal{X}$. As noted in the introduction, these constraints arise as complexity measures on the underlying true data $\mathcal{M}$. One can reliably recover $\mathcal{X}$ using CP algorithms of the type

$$\min h(\mathcal{X}) \quad \text{s.t.} \|\mathcal{Y} - \mathcal{A}(\mathcal{X})\| \le \sigma_n, \tag{2}$$

where $h(\mathcal{X}) \in \mathbb{R}^+ \cup \{0\}$ is a nonnegative real valued mapping that measures the complexity of the true data $\mathcal{X}$ and σ_n^2 is the additive noise variance. Note that equivalently one can also solve for a constrained formulation:

$$\min \|\mathcal{Y} - \mathcal{A}(\mathcal{X})\| \quad \text{s.t.} \, h(\mathcal{X}) \le c_0, \tag{3}$$

for an appropriate choice of the constant c_0.

The t-product

In this section, we review the t-product proposed by Kilmer and Martin (2011b) and further analyzed by Braman (2010) and Kilmer et al. (2013). We will focus on 3D tensors for ease of exposition and interpretation. We begin by viewing a 3D tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ as an $n_1 \times n_2$ matrix of tubes (vectors oriented into the page), whose i,jth entry, denoted by $\mathcal{X}_{i,j}$, is given by $\mathcal{X}_{i,j} \triangleq \mathcal{X}(i,j,:)$. One can consider a $n_1 \times 1 \times n_3$ tensor as a vector of tubes. We will call such tensors *oriented matrices* and denote them by $\vec{\mathbf{M}}$, with the jth tubal element denoted by $\vec{\mathbf{M}}_j$. To define the 3D tensor as a linear operator on the set of oriented matrices $\vec{\mathbf{M}}$, Braman (2010) defines a commutative multiplication operation between two tubes $\vec{\mathbf{v}} \in \mathbb{R}^{1\times 1\times n_3}$ and $\vec{\mathbf{u}} \in \mathbb{R}^{1\times 1\times n_3}$ resulting in another tube of same length. This commutative operation is given by circular convolution denoted by $\star$. Note that $\vec{\mathbf{v}}\star\vec{\mathbf{u}} = \vec{\mathbf{u}}\star\vec{\mathbf{v}}$. Under this multiplication operation and the usual vector addition, the operation of a tensor $\mathcal{X}$ of size $n_1 \times n_2 \times n_3$ on $\vec{\mathbf{M}}$ of size $n_2 \times 1 \times n_3$ is another oriented matrix $\vec{\mathbf{Z}} = \mathcal{X}\star\vec{\mathbf{M}}$ of size $n_1 \times 1 \times n_3$ whose ith tubal element is given by $\vec{\mathbf{Z}}_i = \sum_{j=1}^{n_2} \mathcal{X}_{i,j}\star\vec{\mathbf{M}}_j$. This is illustrated in Figure 1. One can extend this construction and definition to describe the multiplication of two tensors $\mathcal{X}$ and $\mathcal{Y}$ of sizes $n_1 \times n_2 \times n_3$ and $n_2 \times k \times n_3$, respectively, resulting in a tensor $\mathcal{C} = \mathcal{X}\star\mathcal{Y}$ of size $n_1 \times k \times n_3$. This multiplication between the tensors is referred to as the *t-product*.

Tensor singular value decomposition under the tensor product

Using the notion of the t-product and viewing a 3D tensor as a linear operator over the set of oriented matrices, one can compute the tSVD of an $n_1 \times n_2 \times n_3$ tensor $\mathcal{X}$ as shown in Figure 2 and given in the following equation:

$$\mathcal{X} = \mathcal{U}\star\mathcal{S}\star\mathcal{V}^\top, \tag{4}$$

where $\mathcal{U}$ and $\mathcal{S}$ are orthogonal tensors of size $n_1 \times n_1 \times n_3$ and $n_2 \times n_2 \times n_3$, respectively, and $\mathcal{S}$ is a diagonal tensor of size $n_1 \times n_2 \times n_3$. The construction of the tSVD has the same form as the matrix singular value decomposition $\mathbf{X} = \mathbf{USV^T}$ except that t-product and tensor-transpose operations replace the matrix product and matrix transpose. A tensor can be regarded as a matrix of fibers or tubes along the third dimension of a tensor $\mathcal{X}$. The tSVD is analogous to a matrix SVD if we assume that the diagonal tensor $\mathcal{S}$ consists of singular "tubes" or "vectors" on the diagonal, analogous to singular values on the diagonal in traditional SVD. The component tensors $\mathcal{U}$ and $\mathcal{V}$ obey the orthogonality conditions $\mathcal{U}^\top\star\mathcal{U} = \mathcal{I}$ and $\mathcal{V}^\top\star\mathcal{V} = \mathcal{I}$. Definitions for tensor transpose $(\cdot)^\top$ and identity tensor $\mathcal{I}$ are given below.Definition 1.

Tensor transpose: Let $\mathcal{X}$ be a tensor of size $n_1 \times n_2 \times n_3$, then $\mathcal{X}^\top$ is the $n_2 \times n_1 \times n_3$ tensor obtained by transposing each of the frontal slices and then reversing the order of transposed frontal slices two through n_3.Definition 2.

Identity tensor: The identity tensor $\mathcal{I} \in \mathbb{R}^{n\times n\times n_3}$ is a tensor whose first frontal slice is the $n \times n$ identity matrix, and all other frontal slices are zero.

Because the t-product is defined via circulant convolution, the computation of tSVD can be efficiently calculated using the fast Fourier transform (FFT). To compute the tSVD of a 3D tensor $\mathcal{X}$, we first apply the FFT along the third dimension and store the Fourier transformed tensor $\hat{\mathcal{X}}$. We then compute the standard matrix SVD of each frontal slice of the transformed tensor $\hat{\mathcal{X}}$

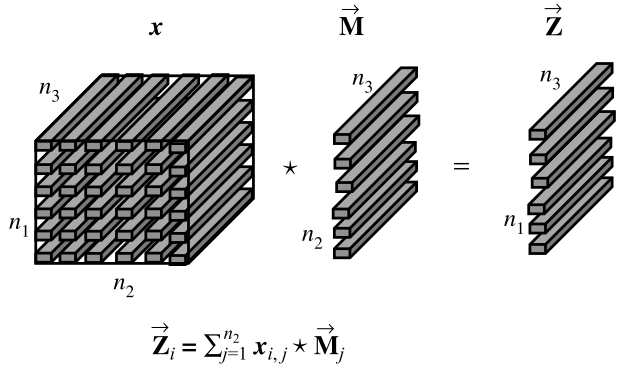

Figure 1. 3D tensors as operators on oriented matrices.

and store a temporary slice of the component tensors $\hat{\mathcal{U}}$, $\hat{\mathcal{S}}$, and $\hat{\mathcal{V}}$. Finally, we apply an inverse FFT to the third dimension of the component tensors to compute the final tSVD decomposition. This concept of the tSVD can be recursively extended to define the t-product and tSVD for higher order tensors (Martin et al., 2013). For example, for 4D data, one can begin by considering it as a matrix of matrices. Then one defines a commutative operation that maps two matrices of same size and outputs another matrix of the same size. This commutative operation is defined to be the 2D circular convolution. Operationally, in the extension of tSVD to higher dimensions, the Fourier transform and inverse Fourier transform are applied recursively in dimension three through p. Note that for our application, $p = 4$ or $p = 5$. Algorithm 1 outlines the steps to compute the tSVD for an order-p tensor. Note that in the algorithm, $\hat{\mathcal{X}}$ is an order p tensor and the notation $\hat{\mathcal{X}}(:,:,i)$ describes the ith frontal slice of size $n_1 \times n_2$ of the tensor $\hat{\mathcal{X}}$ when the tensor is stored in MATLAB format; see Martin et al. (2013) for more details.

The general t-product framework (including computation of tSVD) requires one to fix an orientation of the tensor, i.e., fix the dimensions for the first and second indices. In contrast to the other tensor decomposition methods, which are orientation independent, the fixing of dimension orders makes the tSVD orientation dependent. In many applications where the tensor order, i.e., the number of dimensions, is not very large, such as the one considered in this paper where the tensors are of orders four and five, one can find the best orientation based on some prior experiments with synthetic and real data.

Seismic data completion using tensor singular value decomposition

Based on tSVD, one can define the notion of the tensor rank (Kilmer et al., 2013). The multirank of a tensor using tSVD was defined to be a vector of the ranks of the frontal slices $\hat{\mathcal{X}}(:,:,i), i = 1, 2, \ldots, N$, where $N = n_3 n_4 \ldots n_p$ (for the notation, see Algorithm 1). One can take the sum of the ranks of these frontal slices as a complexity measure. This motivates a CP algorithm for recovering $\mathcal{X}$ from linear measurements by minimizing the sum of the elements of the tubal-rank vector of $\mathcal{X}$ as

$$\text{CP: } \min \sum_{i=1}^{N} \text{rank}(\hat{\mathcal{X}}(:,:,i)) \quad \text{s.t. } \mathcal{Y} = \mathcal{A}(\mathcal{X}). \tag{5}$$

However, as in the recovery of undersampled matrices via linear operators by minimizing rank, the CP problem is known to be computationally infeasible; see Recht et al. (2010). We therefore relax the complexity measure to a norm that we call as the TNN given by $\|\mathcal{X}\|_{\text{TNN}} = \sum_{i=1}^{N} \|\hat{\mathcal{X}}(:,:,i)\|_{\text{nuc}}$, where $\|\cdot\|_{\text{nuc}}$ denotes the Schatten-1 norm on the matrix singular values in the argument (Watson, 1992) also known as the nuclear norm. For the noiseless case, we solve for OPT_TNN(1) and for the noisy case, we solve for OPT_TNN(2) as outlined below:

Algorithm 1. tSVD

Input: $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \ldots \times n_p}$
$N = n_3 n_4 \ldots n_p$
$\hat{\mathcal{X}} = \mathcal{X}$; % Initialization
for $i = 3$ to p **do**
  $\hat{\mathcal{X}} \leftarrow fft(\hat{\mathcal{X}}, [], i)$;
end for
for $i = 1$ to N **do**
  $[\hat{\mathbf{U}}, \hat{\mathbf{S}}, \hat{\mathbf{V}}] = \text{svd}(\hat{\mathcal{X}}(:,:,i))$
  $\hat{\mathcal{U}}(:,:,i) = \hat{\mathbf{U}}; \hat{\mathcal{S}}(:,:,i) = \hat{\mathbf{S}}$
  $\hat{\mathcal{V}}(:,:,i) = \hat{\mathbf{V}}$;
end for
for $i = 3$ to p **do**
  $\mathcal{U} \leftarrow \text{ifft}(\hat{\mathcal{U}}, [], i); \mathcal{S} \leftarrow \text{ifft}(\hat{\mathcal{S}}, [], i)$;
  $\mathcal{V} \leftarrow \text{ifft}(\hat{\mathcal{V}}, [], i)$;
end for

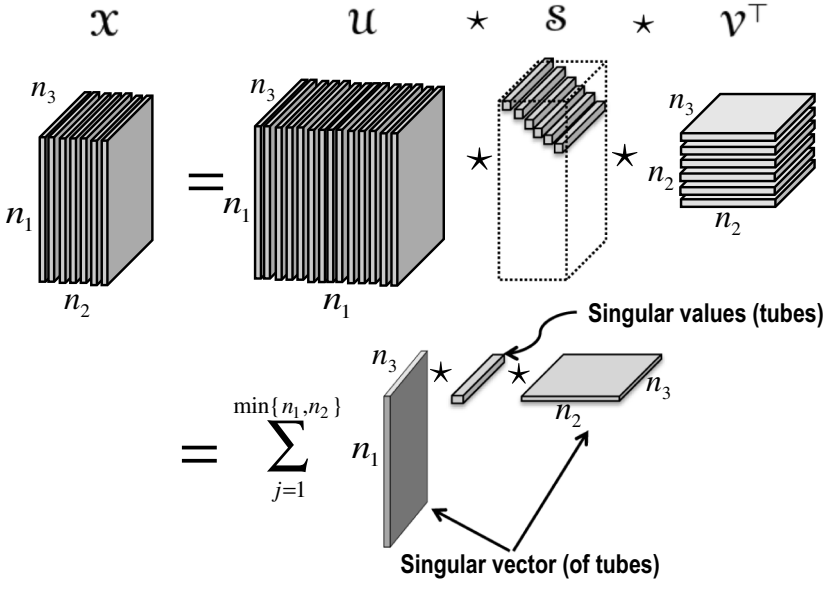

Figure 2. The tSVD of an $n_1 \times n_2 \times n_3$ tensor. The construction of the tSVD is similar to the matrix SVD except that the t-product and tensor transpose substitute the equivalent matrix operations. Like the matrix SVD, the tSVD can also be written as the sum of outer tensor products.

$$\text{OPTTNN(1)}: \min \sum_{i=1}^{N} \|\hat{\mathcal{X}}(:,:,i)\|_{\text{nuc}} \quad \text{s.t.}\, \mathcal{Y} = \mathcal{A}(\mathcal{X}) \tag{6}$$

and

$$\text{OPTTNN(2)}: \min \|\mathcal{Y} - \mathcal{A}(\mathcal{X})\|_2^2 + \lambda \sum_{i=1}^{N} \|\hat{\mathcal{X}}(:,:,i)\|_{\text{nuc}}. \tag{7}$$

Note that although in real seismic applications the noiseless case does not arise, we consider it here for the sake of completeness. In OPT_TNN(2), λ is the low (tensor)-rank tuning factor. The proposed optimization problem is convex (Semerci et al., 2014; Zhang et al., 2014) and, therefore, can be solved using existing techniques.

We would like to point out again that the TNN, as derived from the tSVD, differs significantly from the sum of nuclear norms approach used by Gandy et al. (2011) and Kreimer et al. (2013). tSVD computes TNN by first applying Fourier transform *recursively* along the different dimensions and then taking the sum of nuclear norms of the frontal slices of the resulting tensor, which are $N = n_3 \times n_4 \times n_5$ in number. In contrast, Kreimer et al. (2013) use the nuclear norms of matrices formed by mode unfoldings along the five modes of the tensor if operating in the space-time domain, or along the four modes if operating in the *f-x* domain.

A Kolmogorov Smirnov (KS) test method determines the choice of λ as discussed in the section "Selection of regularization parameter." In addition, the noisy and noiseless versions of the optimization problem can be solved using the alternating direction method of multipliers (ADMM) methods presented in the following section, "Algorithm for solving the optimization problems." These methods are also used by Kreimer et al. (2013) to solve the resulting optimization problem. ADMM methods are ideally suited to the downsampling problem because the ADMM class of algorithms converges in a few number of iterations but requires the typically expensive projection of the residual onto the null space of the forward operator. Note that the projection onto the null space of the downsampling operator is very inexpensive to compute. If, however, the data were acquired using a generic sampling operator, such as in the case in which the traces were randomly summed or subtracted from one another, the projection would be very expensive to compute. In these cases, first-order or incremental methods (Bertsekas, 2011; Juditsky and Nemirovski, 2012) would be better suited. See Ely and Aeron (2013) for an application of such methods to the problem of hydraulic fracture monitoring.

Algorithm for solving the optimization problems

To solve OPT_TNN(1) and OPT_TNN(2), we use the ADMM (Boyd, 2011). The ADMM method, as adapted to our case, works by dividing the variable in two, with one variable capturing the TNN penalty and the other variable capturing the data-matching constraints (as an indicator function). In particular, the ADMM recursion steps (in scaled form) for OPT_TNN(1) are given by

$$\mathcal{X}^{k+1} = \arg\min_{\mathcal{X}} \left\{ \mathbf{1}_{\mathcal{Y}=\mathcal{A}(\mathcal{X})} + \frac{\rho}{2} \|\mathcal{X} - (\mathcal{Z}^k - \mathcal{B}^k\|_F^2 \right\}, \tag{8}$$

$$\mathcal{Z}^{k+1} = \arg\min_{\mathcal{Z}} \left\{ \frac{1}{\rho} \|\mathcal{Z}\|_{\text{TNN}} + \frac{1}{2} \|\mathcal{Z} - (\mathcal{X}^{k+1} + \mathcal{B}^k)\|_F^2 \right\}, \tag{9}$$

and

$$\mathcal{B}^{k+1} = \mathcal{B}^k + (\mathcal{X}^{k+1} - \mathcal{Z}^{k+1}), \tag{10}$$

where $\rho > 0$, $\mathbf{1}_{\mathcal{Y}=\mathcal{A}(\mathcal{X})}$ is the indicator function of the set $\{\mathcal{X}: \mathcal{A}(\mathcal{X}) = \mathcal{Y}\}$ that assumes a value of 0 for all $\mathcal{X}$ satisfying the linear constraint and a value of ∞ otherwise.

The closed-form solution to equation 8, when $\mathcal{A}$ corresponds to the sampling operator, is given by

$$\mathcal{X}^{k+1} = (\mathbf{I} - \mathcal{A})(\mathcal{Z}^k - \mathcal{B}^k) + \mathcal{Y}, \tag{11}$$

Algorithm 2. The ADMM algorithm for seismic data completion using tSVD

$\mathcal{X} = \mathcal{Z} = \mathcal{B} = 0$ // Initialize variables.
$N = n_3 n_4 ... n_p$
while Not Converged **do**
  if Constrained **then**
    // *Projection onto* $\mathcal{Y} = \mathcal{A}(\mathcal{X})$ when $\mathcal{A}$ is the sampling operator
    $\mathcal{X} = (\mathbf{I} - \mathcal{A})(\mathcal{Z} - \mathcal{B}) + \mathcal{Y}$
  else
    $\mathcal{X} = (\rho\mathbf{I} + \mathcal{A})^{-1}(\rho(\mathcal{Z} - \mathcal{B}) + \mathcal{Y})$;
  end if
  $\mathcal{M} = \mathcal{X} + \mathcal{B}$
  // *Calculate tSVD and shrink in the Fourier domain*
  $\mathcal{D} = \mathcal{M}$
  for $i = 3$ to N **do**
    $\mathcal{D} \leftarrow \text{fft}(\mathcal{D}, [], i)$;
  end for
  // *shrinking via singular value thresholding on each slice*
  for $i = 1$ to N **do**
    $[\mathbf{U}, \mathbf{S}, \mathbf{V}] = (\mathcal{D}(:,:,i))$;
    // *If constrained* $\varepsilon = \frac{1}{\rho}$, *If unconstrained* $\varepsilon = \frac{\lambda}{\rho}$
    $\mathbf{S} = Sh_{\varepsilon}[\mathbf{S}]$;
    $\mathcal{U}(:,:,i) = \mathbf{U}; \mathcal{S}(:,:,i) = \mathbf{S}; \mathcal{V}(:,:,i) = \mathbf{V}$;
  end for
  // *Going back to the original domain*
  for $i = 3$ to p **do**
    $\mathcal{U} \leftarrow \text{ifft}(\mathcal{U}, [], i); \mathcal{S} \leftarrow \text{ifft}(\mathcal{S}, [], i); \mathcal{V} \leftarrow \text{ifft}(\mathcal{V}, [], i)$;
  end for
  $\mathcal{Z} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$
  // *Dual Update*
  $\mathcal{B} = \mathcal{B} + \mathcal{X} - \mathcal{Z}$
end while

where $\mathbf{I}$ denotes the identity operator. For the unconstrained optimization case OPT_TNN(2), the update equation for $\mathcal{X}^{k+1}$ becomes

$$\mathcal{X}^{k+1} = (\rho\mathbf{I} + \mathcal{A})^{-1}(\mathcal{Y} + \rho(\mathcal{Z}^k - \mathcal{B}^k)). \qquad (12)$$

The solution to equation 9 is obtained by singular value thresholding of the frontal slices in the Fourier domain as shown in Algorithm 2, which shows the overall pseudocode for solving the problems OPT_TNN(1) and OPT_TNN(2). In Algorithm 2, Sh_ϵ is an elementwise shrinkage function that applies a soft thresholding according to equation 11 as

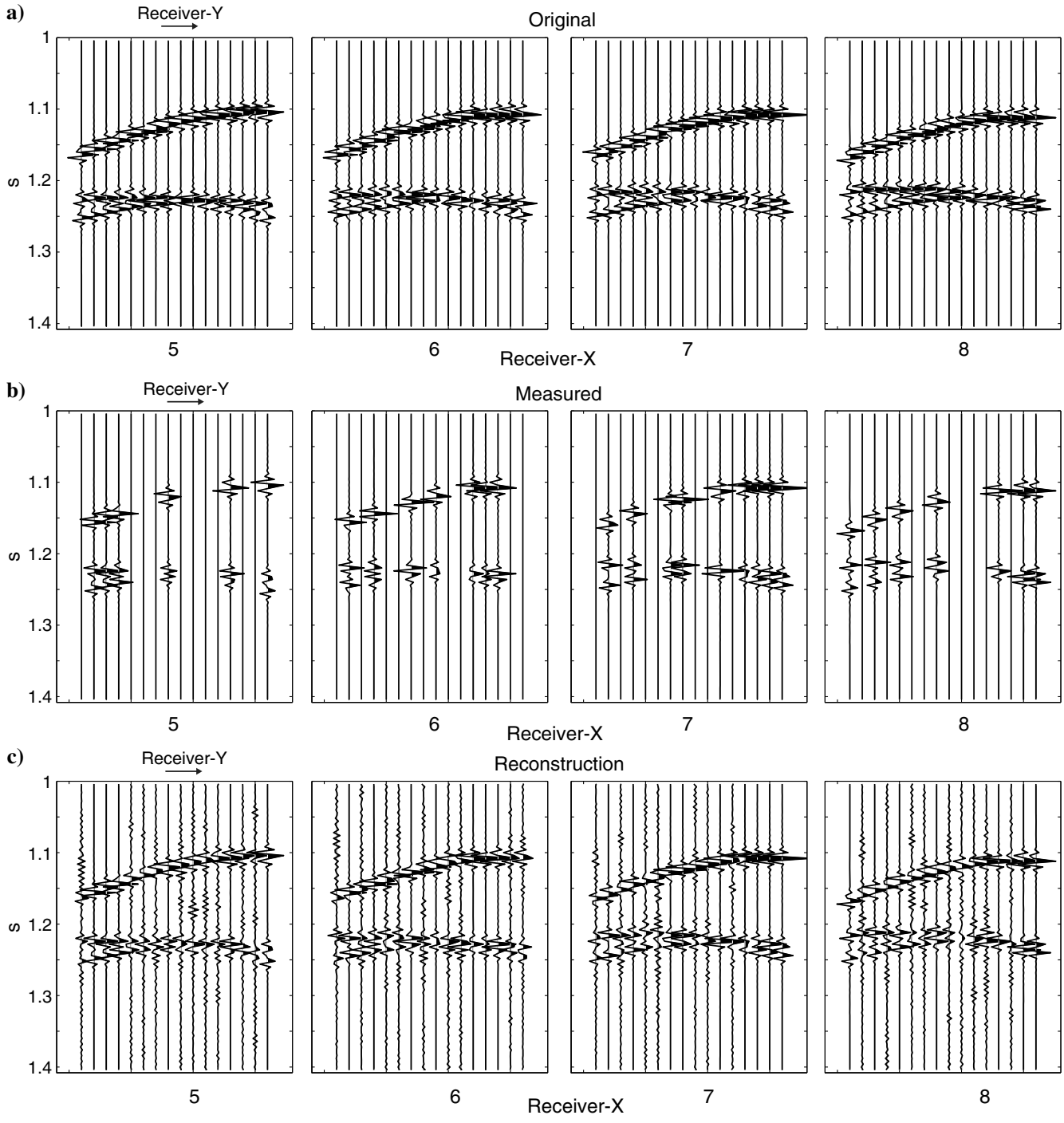

Figure 3. (a) This figure shows the full synthetic data for four different receiver-source slices as well as (b) the undersampled measured data for the case in which 60% of the traces were removed. In addition, (c) the reconstruction for the 5D slices is shown as well. All subfigures are shown with the same gain and clip values.

$$Sh_\varepsilon[x] = \begin{cases} x-\varepsilon, & \text{if } x > \varepsilon, \\ x+\varepsilon, & \text{if } x < \varepsilon, \\ 0, & \text{otherwise.} \end{cases} \quad (13)$$

For the constrained noiseless case $\varepsilon = (1/\rho)$ and for the unconstrained noisy case $\varepsilon = (\lambda/\rho)$ where ρ is the step size and λ is the low-rank tuning factor controlling the severity of the TNN penalty. The convergence criteria in Algorithm 2 are set according to relative tolerance. For all of the instances of the algorithm, we set ρ to be equal to 1. The algorithm is robust for the choice of this parameter, although for some very small or large values, the performance of the algorithm degrades.

PERFORMANCE EVALUATION: SYNTHETIC DATA

Compressibility of seismic data in the tensor singular value decomposition domain

To demonstrate the compressibility of the seismic data with nonlinear reflectors, we generate a synthetic 5D survey in which sources and receivers were placed on a 16×16 grid with 50-m shot and receiver spacing in the x- and y-directions. Three synthetic Born scatters were placed below the surface. Traces consisting of N_t samples were generated for all the possible source-receiver combinations, resulting in a 5D tensor with dimensions $16 \times 16 \times 16 \times 16 \times N_t$. Figure 3 shows several receiver gathers from the 5D synthetic survey for the source located at indices $s_x = 3$ and $s_y = 7$ with each subplot showing four receiver gathers. Each of the four subplots contains gathers for the constant X receiver indices (5,6,7,8) and all receiver Y indices. The tSVD was applied to th[pe synthetic data. Figure 4 shows the decay of the singular values of the matrices $\hat{\mathcal{X}}(:,:,i)$ as computed using the tSVD. For the synthetic data case, they obey an empirical power law decay, which implies that $\sum_{i=1}^{N} \|\hat{\mathcal{X}}(:,:,i)\|_{\text{nuc}}$ is a good complexity measure in the tSVD domain. As a result, we expect Algorithm 2 to recover the data from a limited number of measurements.

Simple reflectors

To evaluate our algorithm, we use the synthetic data set described previously and remove 10% to 90% of the measured traces at 5% intervals, resulting in highly undersampled data as shown in Figure 3b for 60% of traces removed. Algorithm 2 was then applied in two ways: (1) Recovering the data by completion of 4D tensors frequency by frequency and (2) recovering the entire 5D seismic volume at once. In addition to testing a random downsampling operator, we experiment with a regularly spaced operator in which every second or third shot or receiver was removed. Under the regular downsampling, completion using the TNN proved to be unreliable when only 50% of the traces were removed. The reason that regular downsampling does not work well can be attributed to the reasons why it fails in the case of simple matrix completion. One reason behind this is that for regular downsampling, the resulting matrix (padded with zeros at places in which the data are not sampled) can be of the same or lower rank as the original matrix. Clearly in this case, the nuclear-norm heuristic will fail. On the other hand, such a case does not happen with high prob-

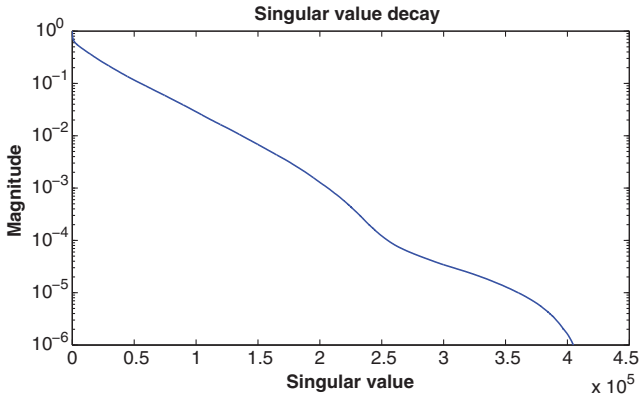

Figure 4. Decay of singular values of the synthetic seismic data that empirically obey a power law decay.

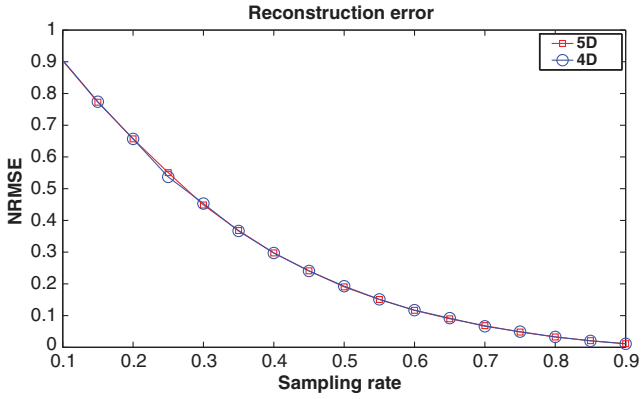

Figure 5. Reconstruction error as function of the sampling fraction for the 4D frequency by frequency and full 5D reconstruction. Reconstruction frequency by frequency results in nearly identical performance as completion in the original 5D domain.

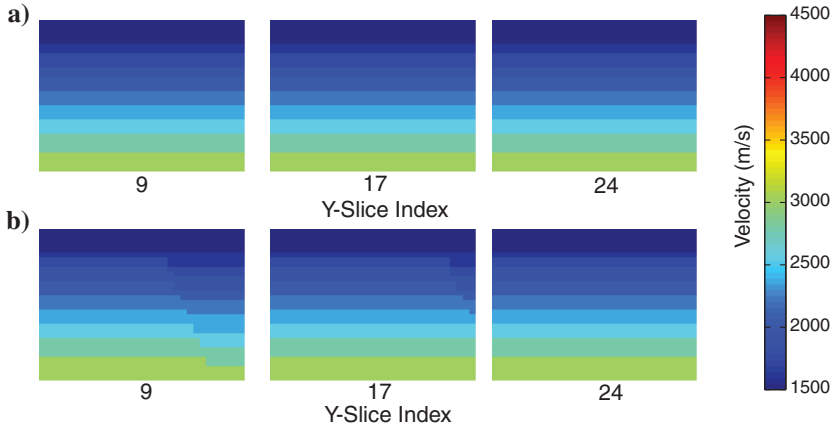

Figure 6. Three vertical slices of the velocity model used to generate the synthetic shots. (a) The simplest velocity model with water and sedimentary layers. (b) Flat layered velocity model with a normal fault.

ability when using random downsampling. We refer the interested reader to Recht et al. (2011) for a more technical discussion, implicit in the analysis there. An analysis of such conditions for our case is a subject of current investigations. The normalized root-mean-square error (NRMSE) measures the recovery error. For the synthetic data, Figure 5 shows the performance as function of sampling rate. Although the NRMSE is significant for highly undersampled data, the reflectors are still visible in the reconstruction; refer to Figure 3. Therefore, the NRMSE appears to underestimate reconstruction quality. For all sampling rates, the completion performance in the 5D or 4D domain is nearly identical, with reconstruction error varying less than 1% as shown in Figure 5.

Although many reconstruction techniques must operate in the Fourier domain to exploit the sparsity of the data cube in the *f-x* domain, the performance of our method does not appear to depend on the temporal Fourier transform. This is due to the fact that the construction of the tSVD involves the recursive application of the Fourier transform along the temporal as well as the spatial dimensions of tensor. After taking the Fourier transform along the temporal dimensions, our method implicitly operates in the spatial Fourier domain. Given that time and Fourier domain reconstructions of the tensor are nearly identical, it could be computationally advantageous to reconstruct the data in a frequency-by-frequency manner because this is trivial to parallelize across numerous nodes. For a 5D data set with N_t samples, the completion would only need to be performed on $N_t/2$ 4D tensors due to the symmetry of the Fourier transform. The number of computations could further be reduced if the signal is band limited and only a fraction of the $N_t/2$ frequencies is needed. Given that the tSVD shrinkage operator already acts in the frequency domain and the SVD is already computed on a complex matrix, transforming the data into the Fourier domain incurs no additional computational costs.

Despite these advantages of the Fourier domain, we choose to work in the time domain for most of the synthetic and real data sets to simplify the approach and to reduce the number of regularization parameters. In particular, if the reconstruction algorithm was applied to noisy data in the *f-x* domain, then a regularization parameter would need to be selected for each frequency. This is due to the fact that the source wavelet is band limited, resulting in a different S/N across frequencies, which in turn necessitates optimizing over the regularization parameter for each frequency.

Finite-difference synthetic

Because Born scatters inaccurately represent typical data sets, we generate a more realistic set of synthetic survey over two velocity models of varying complexity: (1) a flat layered model and (2) a layered model with a normal fault. Figure 6 shows cross sections for the two 2000-m cubed velocity models. For each of the two models, two orthogonal surveys were generated with different shot geometries with the same total number of shots and receivers. Synthetic shots were fired over a 16×16 grid and a 32×32-receiver grid. For the small spatial survey geometry, shots and receivers were placed over a 1000×1000-m area centered in the middle of the velocity model, resulting in shot spacing of 62.5 m and receiver spacing of 31.25 m. For the large spatial geometry, the shots and receivers were spread across the entire 2000×2000-m surface with shot spacing of 125 m and receiver spacing of 62.5 m. We then used PySit, a seismic inversion and forward-modeling toolkit, to generate the synthetic traces using a finite-difference solver such that the resulting data sets had a spatial dimension of $16 \times 16 \times 32 \times 32$ in the shot-receiver domain. For each of the two velocity models and two receiver geometries, we apply Algorithm 1 to calculate the tSVD decomposition of the four data sets. Using these decompositions, we generate truncated rank (compressed) versions of the data sets and compare the error between the low-rank reconstruction and original data as a function of the total number of singular vector pairs used. Figure 7 shows the normalized mean squared error (NMSE) of the reconstructions versus the fraction of singular vector pairs for the wavefield arriving after the direct arrival. From these curves, we see that the survey geometries and complexity of the velocity significantly affect the data's compressibility, with survey geometry being most dramatic. With a wide shot and receiver spacing, each trace contains more unique information than the small spatial geometry, in which each trace contains more redundant information. For any survey geometry, the complexity of the model inversely correlates with compressibility. Given that the more complex velocity model would involve a more complex wavefield geometry, we conclude that more singular vectors would be needed to represent the wavefield. From the demonstrated relationship among compressibility, survey geometry, and the velocity model, we expect a similar behavior for reconstruction of missing traces. Performance could also be improved by applying the reconstruction on a small spatiotemporal window to reduce the complexity and nonlinearity of the traces.

For the four surveys, between 25% and 97.5% of the source-receiver pairs are randomly removed, and we use the full noiseless 5D completion algorithm to reconstruct the missing traces. We apply Algorithm 2 to the decimated traces for 500 iterations from the four surveys and different sampling rates. Figure 7 shows the NMSE for the wavefield after the direct arrival. Regardless of

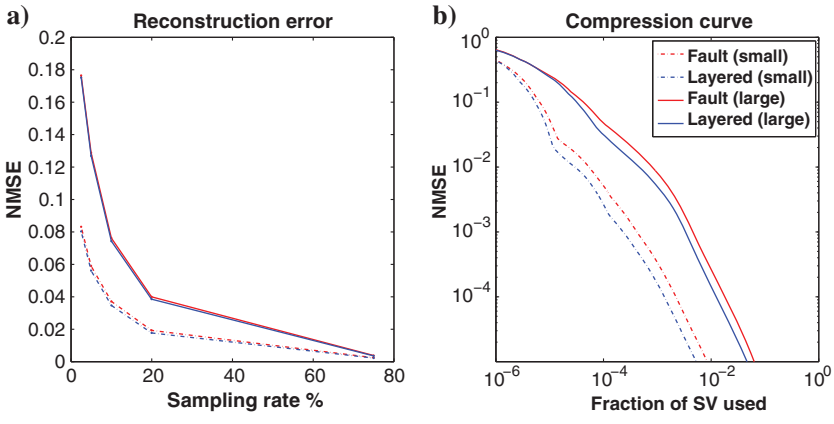

Figure 7. (a) Reconstruction error for the four data sets as a function of the downsampling rate. (b) Compressibility of the data sets in the tSVD domain as a function of the fraction of singular vector pairs used to reconstruct the data set. Note that if the data are more compressible in the tSVD domain, then the corresponding recovery error is also small.

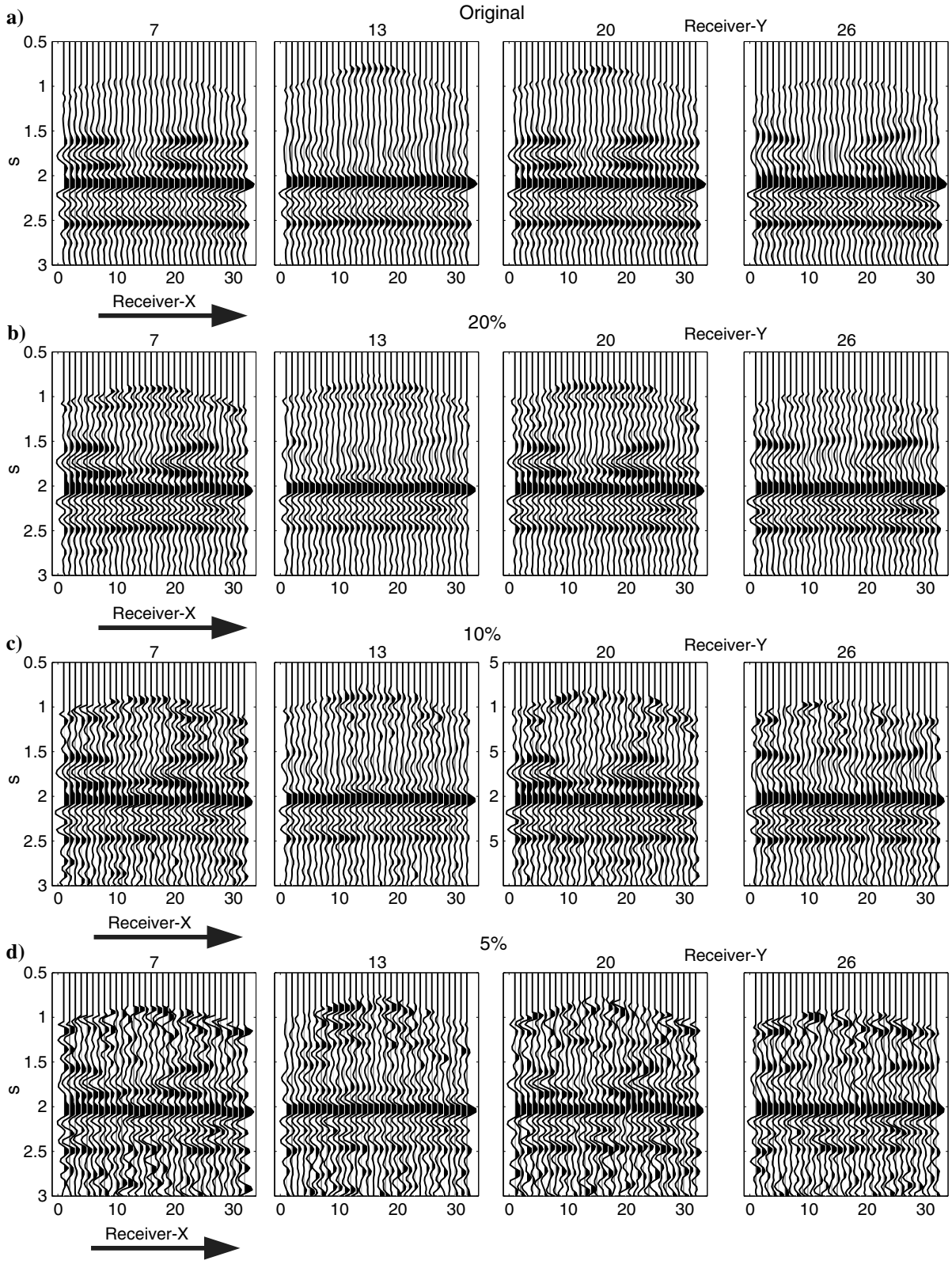

Figure 8. The top set of traces in this figure shows four gathers for a shot at the center of the survey at (x, y) source indices of (8,8). Each panel shows a constant receiver-y-index. The bottom three panels show the reconstructed traces at three different sampling rates, 20%, 10%, and 5%.

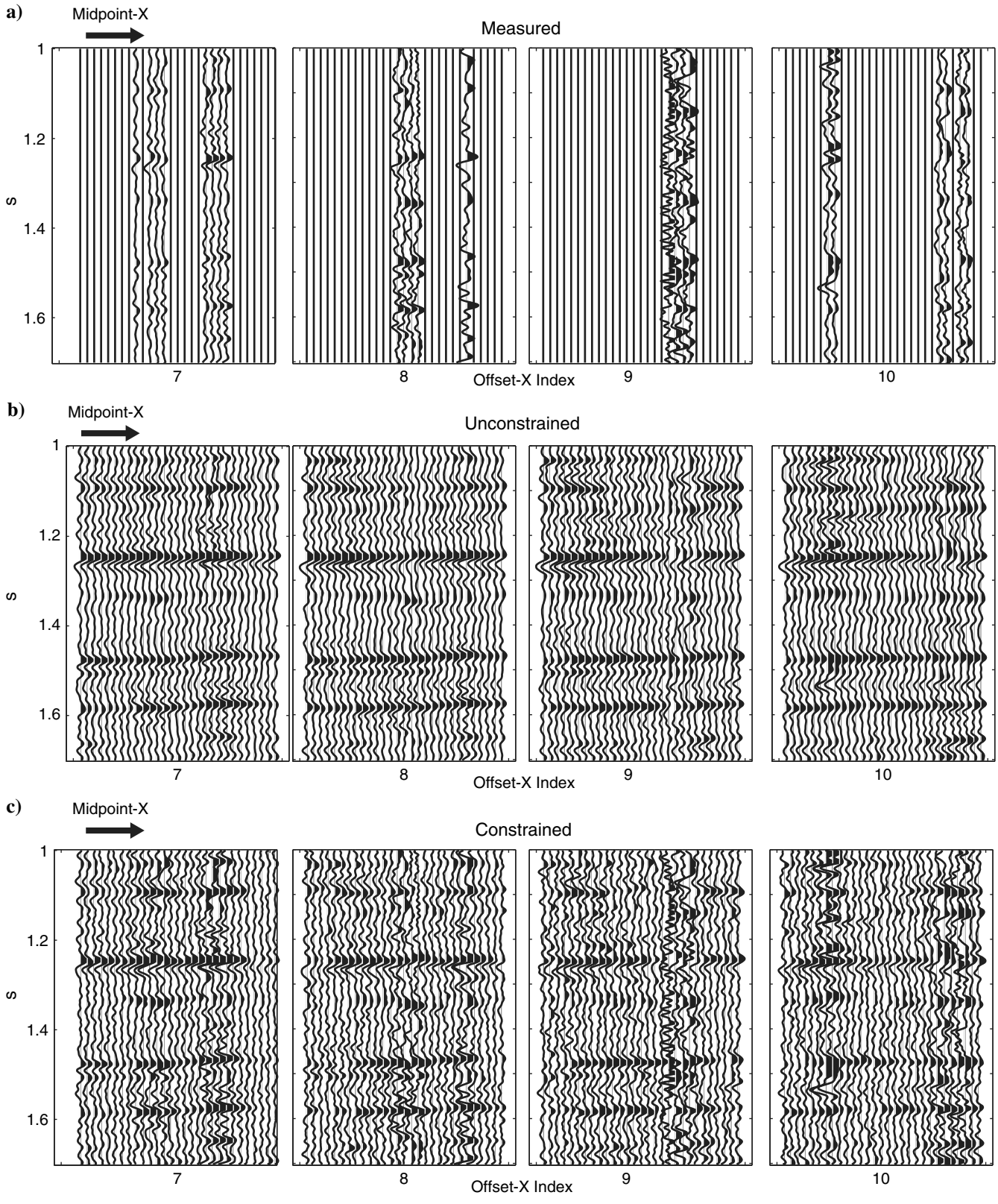

Figure 9. (a) The original sparsely sampled field data from the Western Canadian Sedimentary Basin for a fixed y-midpoint of 11 and a y-offset of 6. (b) The denoised and reconstructed traces using OPT_TNN(2). (c) The results of applying OPT_TNN(1), reconstruction only, to the original traces. Many of the reflectors are better reconstructed using OPT_TNN(2) because the constrained optimization over fits the noise.

the velocity model, the reconstruction error behaved comparably across the four cases and the error appears to be minimal for down sampling rates of 10% or less. Figure 8 shows the original and reconstructed traces from model B and small spatial survey geometry for a single slice at several downsampling rates. At downsampling rates lower than 10%, many of the reflections in the wavefield become severely distorted and difficult to discern. For ease of viewing the relevant reflections, we null the traces at the direct arrival.

Although the type of down sampling in which we randomly remove source-receiver pairs would not be encountered in a real survey, the experiment provides a theoretical method of validating the reconstruction technique. Normally, the data would instead be completed in other domains, such as the midpoint offset domain. To generate full 5D finite-difference data in the midpoint offset domain without binning, a forward solver would be needed for each trace instead of each source location as in the source receiver domain. Generating data in this manner would make it computationally burdensome to create true data to test the completion algorithm. In the following section, we test the algorithm in the midpoint offset domain for real data.

PERFORMANCE EVALUATION ON FIELD DATA

In addition to synthetic data, we test the algorithm on a survey from the Western Canadian Sedimentary Basin. This is a gridded data set consisting of 29 midpoints and 12 offsets in the x- and y-directions, respectively, or a volume with spatial dimensions of $29 \times 29 \times 12 \times 12$. The midpoint grid spacing is 26 and 52 m for the x- and y-directions with offset spacing from 0 to 1400 m. Roughly 121,000 midpoint and offset combinations are possible; however, only 16,060 traces were recorded, resulting in a sampling rate of approximately 8%. Figure 9 top shows the original reflection data from the survey for several x-offsets with a fixed y-midpoint of 11 and a y-offset of 6 in which numerous reflectors are present.

To reconstruct the traces, we solve the constrained and unconstrained optimization with a λ of 0.0038, determined by the KS test. Figure 9b and 9c shows the reconstructed results for a fixed y-midpoint of 11 and y-offset of 6. This volume coincides with the one used by Kreimer and Sacchi (2012a), who use HOSVD to reconstruct the same data set. We observe in the figure that the two completion algorithms can reconstruct the reflectors, with the noisy reconstruction having noticeably better performance. For all of these gathers, the noisy optimization scheme leads to better reconstruction of the missing seismic data. Because the constrained optimization assumes that there is little or no noise in the measured data, the algorithm makes the noise look coherent, resulting in a poor reconstruction.

To test the impact of our completion algorithm and confirm whether it is removing signal, we stack the data. Figure 10 shows a single slice of the stacked image for a constant common midpoint x of 15 for the two reconstructions and the stack of the input traces. Both reconstructions have significantly less noise than the raw stacked image but have little difference in quality between them. Given that the noisy optimization had better completion, we would expect it to give improved results for other postprocessing techniques.

Selection of regularization parameter

In the presence of noise, the minimization problem OPT_TNN(2) introduces the unknown parameter λ to determine the trade-off between measurement error and the TNN of $\boldsymbol{\mathcal{X}}$. The regularization term λ is typically determined through heuristics such as the L-curve method (Hansen and O'Leary, 1993). However, choosing a suitable λ based on the L-curve can be difficult to determine and challenging to automate. To determine λ, we use a variation in the KS test, described in detail by Aeron et al. (2011). The method computes the residual for the entire range of the regularization parameters. From these sets of residuals, we compute the KS statistic and P-value between each distribution and the two distributions with the smallest

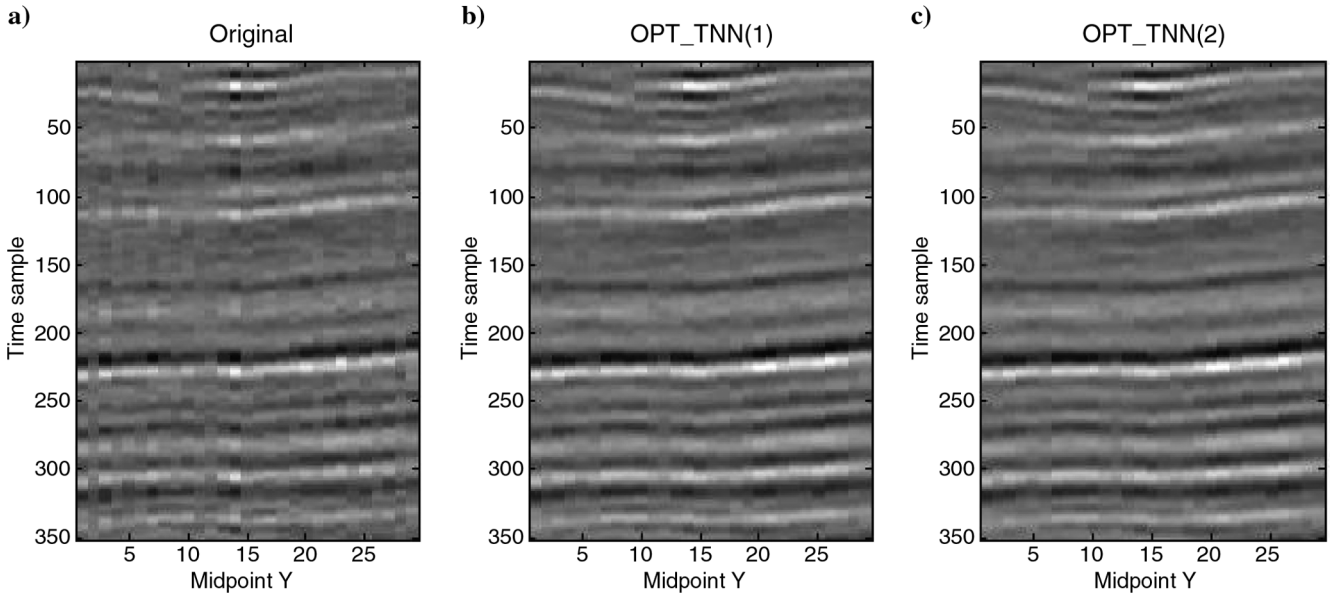

Figure 10. (a) Stack of the original data set. (b) Stack of the reconstructed data with no-noise correction. (c) Stack of the joint noise removal and reconstruction.

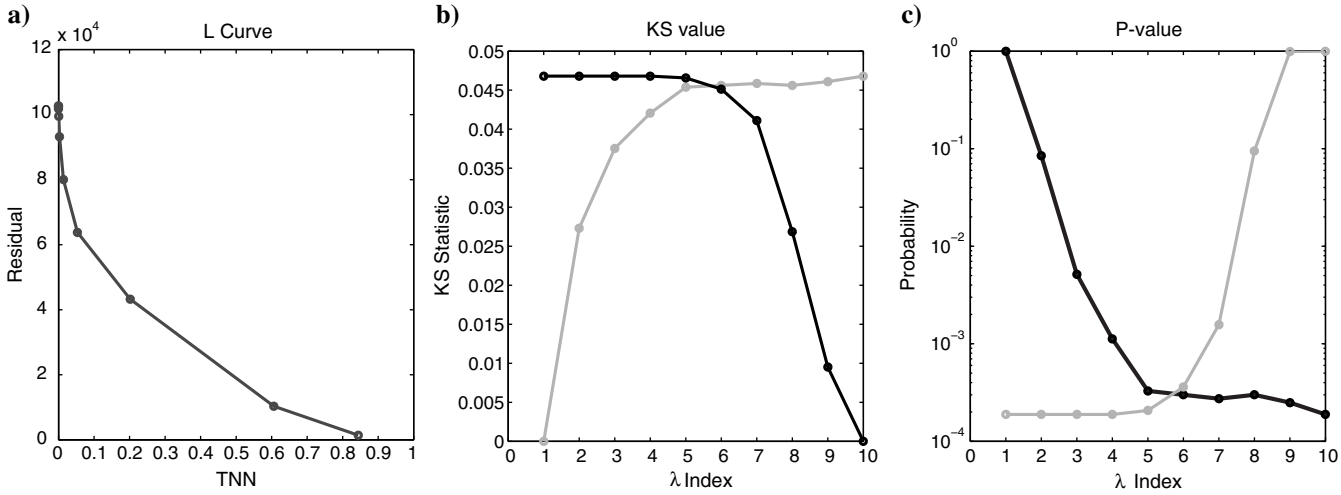

Figure 11. (a) The L-curve, residual error as function of TNN of 10 solutions. (b and c) KS statistic and P-values for the range of λs. The P and KS curves show that between the sixth or seventh value of λ is optimal, a value of roughly 0.0038.

and largest regularization parameters to generate two curves of the KS statistic. The optimal operating point is the intersection of the these two curves, Figure 11b and 11c.

In this particular experiment, we generate simulations for 10 different values logarithmically spaced from 10^{-6} to $10^{.45}$. The L-curve method generates a curve having no distinct knee, making it impossible to choose a viable λ, see Figure 11a. The KS test was performed for the 10 values of λ with the residual randomly downsampled by a factor of 10^4 to reduce computational time. The intersection of the KS and P-value gives the same selection of the parameter λ. The KS test resulted in $\lambda = .0038$, which was then used to generate all the plots showing reconstruction of noisy field data.

CONCLUSIONS

We present a novel method for the reconstruction and denoising of incomplete seismic data through the use of a tSVD rank minimization algorithm, and we apply the algorithm to a variety of synthetic data sets and a real data set. We test the algorithm in the *f-x* and the *t-x* domain. In addition, we demonstrate how the survey geometry and velocity model complexity can impact the compression and reconstruction of seismic data. Wide receiver shot spacing and a more complex velocity model lead to poor reconstruction, and interpolation is most feasible for tight shot-receiver spacing. Subsequently, we demonstrate that our algorithm can be successfully applied to field data to improve the stacked image.

ACKNOWLEDGMENTS

We would like to thank CGG, Canada, and M. Sacchi from the University of Alberta, Canada, for providing us with the field data. In addition, we would like to thank R. Hewett for his assistance with PySIT and the generation of synthetic data sets. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under grant no. DGE-0806676 and the Computing and Communication Foundations under grant no. NSF-1319653.

REFERENCES

Aeron, S., S. Bose, H.-P. Valero, and V. Saligrama, 2011, Broadband dispersion extraction using simultaneous sparse penalization: IEEE Transactions on Signal Processing, **59**, 4821–4837, doi: 10.1109/TSP.2011.2160632.

Bertsekas, D. P., 2011, Incremental gradient, subgradient, and proximal methods for convex optimization: A survey, *in* S. Sra, S. Nowozin, and S. J. Wright, eds., Optimization for machine learning: MIT Press, 85–115.

Boyd, S., 2011, Distributed optimization and statistical learning via the alternating direction method of multipliers: Foundations and Trends® in Machine Learning, **3**, 1–122, doi: 10.1561/2200000016.

Braman, K., 2010, Third-order tensors as linear operators on a space of matrices: Linear Algebra and Its Applications, 1241–1253, doi: 10.1016/j.laa.2010.05.025.

Curry, W., 2010, Interpolation with Fourier-radial adaptive thresholding: Geophysics, **75**, no. 6, WB95–WB102.

Duijndam, A., M. Schonewille, and C. Hindriks, 1999, Reconstruction of bandlimited signals, irregularly sampled along one spatial direction: Geophysics, **64**, 524–538, doi: 10.1190/1.1444559.

Ely, G., and S. Aeron, 2013, Methods for large scale hydraulic fracture monitoring, *in* IEEE 5th International Workshop Computational Advances in Multi-Sensor Adaptive Processing: IEEE, 272–275.

Ely, G., S. Aeron, N. Hao, and M. Kilmer, 2013, 5D and 4D pre-stack seismic data completion using tensor nuclear norm (TNN): 83rd Annual International Meeting, SEG, Expanded Abstracts, 3639–3644.

Fomel, S., 2003, Seismic reflection data interpolation with differential offset and shot continuation: Geophysics, **68**, 733–744, doi: 10.1190/1.1567243.

Fomel, S., and Y. Liu, 2010, Seislet transform and seislet frame: Geophysics, **75**, no. 3, V25–V38, doi: 10.1190/1.3380591.

Gandy, S., B. Recht, and I. Yamada, 2011, Tensor completion and low-n-rank tensor recovery via convex optimization: Inverse Problems, **27**, 025010, doi: 10.1088/0266-5611/27/2/025010.

Gao, J., M. D. Sacchi, and X. Chen, 2011, A fast rank reduction method for the reconstruction of 5D seismic volumes: 81st Annual International Meeting, SEG, Expanded Abstracts, 3622–3627.

Gao, J., M. D. Sacchi, and X. Chen, 2013, A fast reduced-rank interpolation method for prestack seismic volumes that depend on four spatial dimensions: Geophysics, **78**, no. 1, V21–V30, doi: 10.1190/geo2012-0038.1.

Grasedyck, L., 2010, Hierarchical singular value decomposition of tensors: SIAM Journal on Matrix Analysis and Applications, **31**, 2029–2054, doi: 10.1137/090764189.

Gulunay, N., 2003, Seismic trace interpolation in the Fourier transform domain: Geophysics, **68**, 355–369, doi: 10.1190/1.1543221.

Hansen, P. C., and D. P. O'Leary, 1993, The use of the L-curve in the regularization of discrete ill-posed problems: SIAM Journal on Scientific Computing, **14**, 1487–1503, doi: 10.1137/0914086.

Hennenfent, G., and F. Herrmann, 2006, Seismic denoising with nonuniformly sampled curvelets: Computing in Science Engineering, **8**, 16–25, doi: 10.1109/MCSE.2006.49.

Herrmann, F. J., and G. Hennenfent, 2008, Non-parametric seismic data recovery with curvelet frames: Geophysical Journal International, **173**, 233–248, doi: 10.1111/j.1365-246X.2007.03698.x.
Hindriks, K., and A. Duijndam, 2000, Reconstruction of 3-D seismic signals irregularly sampled along two spatial coordinates: Geophysics, **65**, 253–263, doi: 10.1190/1.1444716.
Hunt, L., J. Downton, S. Reynolds, S. Hadley, D. Trad, and M. Hadley, 2010, The effect of interpolation on imaging and AVO: A Viking case study: Geophysics, **75**, no. 6, WB265–WB274, doi: 10.1190/1.3475390.
Juditsky, A., and A. Nemirovski, 2012, First-order methods for nonsmooth convex large-scale optimization Part II — Utilizing problem's structure: Optimization for machine learning, 149–183.
Kabir, M. N., and D. Verschuur, 1995, Restoration of missing offsets by parabolic Radon transform: Geophysical Prospecting, **43**, 347–368, doi: 10.1111/j.1365-2478.1995.tb00257.x.
Kaplan, S., M. Naghizadeh, and M. Sacchi, 2010, Two dimensional shot-profile migration data reconstruction: 80th Annual International Meeting, SEG, Expanded Abstracts, 3645–3649.
Kilmer, M., K. Braman, N. Hao, and R. Hoover, 2013, Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging: SIAM Journal on Matrix Analysis and Applications, **34**, 148–172, doi: 10.1137/110837711.
Kilmer, M. E., and C. D. Martin, 2011a, Factorization strategies for third-order tensors: Linear Algebra and Its Applications, **435**, 641–658 doi: 10.1016/j.laa.2010.09.020.
Kilmer, M. E., and C. D. Martin, 2011b, Factorization strategies for third-order tensors: Linear Algebra and Its Applications, Special Issue in Honor of G. W. Stewart's 70th birthday, **435**, 641–658, doi: 10.1016/j.laa.2010.09.020.
Kolda, T. G., and B. W. Bader, 2009, Tensor decompositions and applications: SIAM Review, **51**, 455–500, doi: 10.1137/07070111X.
Kreimer, N., and M. Sacchi, 2012a, Reconstruction of seismic data via tensor completion, Statistical Signal Processing Workshop, IEEE, 29–32.
Kreimer, N., and M. Sacchi, 2012b, A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation: Geophysics, **77**, no. 3, V113–V122, doi: 10.1190/geo2011-0399.1.
Kreimer, N., A. Stanton, and M. D. Sacchi, 2013, Tensor completion based on nuclear norm minimization for 5D seismic data reconstruction: Geophysics, **78**, no. 6, V273–V284, doi: 10.1190/geo2013-0022.1.
Kumar, R., A. Aravkin, H. Mansour, B. Recht, and F. Herrmann, 2013, Seismic data interpolation and denoising using SVD-free low-rank matrix factorization: Presented at the 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013.
Lathauwer, L. D., B. D. Moor, and J. Vandewalle, 2000, A multilinear singular value decomposition: SIAM Journal of Matrix Analysis and Applications, **21**, 1253–1278.
Liu, B., and M. D. Sacchi, 2004, Minimum weighted norm interpolation of seismic records: Geophysics, **69**, 1560–1568, doi: 10.1190/1.1836829.
Martin, C. D., R. Shafer, and B. LaRue, 2013, An order-P tensor factorization with applications in imaging: SIAM Journal on Scientific Computing, **35**, A474–A490, doi: 10.1137/110841229.
Naghizadeh, M., and M. Sacchi, 2010a, Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data: Geophysics, **75**, no. 6, WB189–WB202, doi: 10.1190/1.3509468.
Naghizadeh, M., and M. Sacchi, 2010b, Seismic data reconstruction using multidimensional prediction filters: Geophysical Prospecting, **58**, 157–173, doi: 10.1111/j.1365-2478.2009.00805.x.
Naghizadeh, M., and M. D. Sacchi, 2007, Multistep autoregressive reconstruction of seismic records: Geophysics, **72**, no. 6, V111–V118, doi: 10.1190/1.2771685.
Oropeza, V., and M. Sacchi, 2010, A randomized SVD for multichannel singular spectrum analysis (MSSA) noise attenuation, 80th Annual International Meeting, SEG, Expanded Abstracts, 3539–3544.
Oropeza, V., and M. Sacchi, 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: Geophysics, **76**, no. 3, V25–V32, doi: 10.1190/1.3552706.
Recht, B., M. Fazel, and P. A. Parrilo, 2010, Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization: SIAM Review, **52**, 471–501, doi: 10.1137/070697835.
Recht, B., W. Xu, and B. Hassibi, 2011, Null space conditions and thresholds for rank minimization: Mathematical Programming, **127**, 175–202.
Ronen, J., 1987, Wave equation trace interpolation: Geophysics, **52**, 973–984, doi: 10.1190/1.1442366.
Sacchi, M. D., and B. Liu, 2005, Minimum weighted norm wavefield reconstruction for AVA imaging: Geophysical Prospecting, **53**, 787–801, doi: 10.1111/j.1365-2478.2005.00503.x.
Semerci, O., N. Hao, M. Kilmer, and E. Miller, 2014, Tensor-based formulation and nuclear norm regularization for multienergy computed tomography: IEEE Transactions on Image Processing, **23**, 1678–1693, doi: 10.1109/TIP.2014.2305840.
Silva, C. D., and F. J. Herrmann, 2013, Hierarchical Tucker tensor optimization — Applications to 4D seismic data interpolation: Presented at the 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013.
Silva, V. D., and L.-H. Lim, 2008, Tensor rank and the ill-posedness of the best low-rank approximation problem: SIAM Journal on Matrix Analysis and Applications, **30**, 1084–1127, doi: 10.1137/06066518X.
Spitz, S., 1991, Seismic trace interpolation in the *f-x* domain: Geophysics, **56**, 785–794, doi: 10.1190/1.1443096.
Stanton, A., and M. D. Sacchi, 2013, All roads lead to Rome: Predictability, sparsity, rank and pre-stack seismic data reconstruction: CSEG Recorder, **38**.
Stolt, R., 2002, Seismic data mapping and reconstruction: Geophysics, **67**, 890–908, doi: 10.1190/1.1484532.
Trad, D., 2009, Five-dimensional interpolation: Recovering from acquisition constraints: Geophysics, **74**, no. 6, V123–V132, doi: 10.1190/1.3245216.
Trickett, S., L. Burroughs, and A. Milton, 2013, 704, Interpolation using Hankel tensor completion: 83rd Annual International Meeting, SEG, Expanded Abstracts, 3634–3638.
Trickett, S., L. Burroughs, A. Milton, L. Walton, and R. Dack, 2010, Rank-reduction-based trace interpolation: 80th Annual International Meeting, SEG, Expanded Abstracts, 3829–3833.
Watson, G., 1992, Characterization of the subdifferential of some matrix norms: Linear Algebra and Its Applications, **170**, 33–45, doi: 10.1016/0024-3795(92)90407-2.
Xu, S., Y. Zhang, D. Pham, and G. Lambar, 2005, Antileakage Fourier transform for seismic data regularization: Geophysics, **70**, no. 4, V87–V95, doi: 10.1190/1.1993713.
Zhang, Z., G. Ely, S. Aeron, N. Hao, and M. Kilmer, 2014, Novel methods for multilinear data completion and de-noising based on tensor-SVD: *in* IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 3842–3849.
Zwartjes, P., and M. Sacchi, 2007, Fourier reconstruction of nonuniformly sampled, aliased seismic data: Geophysics, **72**, no. 1, V21–V32, doi: 10.1190/1.2399442.