Improving seismic data completion via low-rank tensor optimization

Jonathan Popa, Susan E. Minkoff and Yifei Lou, University of Texas at Dallas*

SUMMARY

The data completion problem involves recovering missing seismic data from incomplete observations. 3D seismic data can be naturally arranged into an order-3 tensor with dimensions for time, offset, and gather number. Redundancy in the data leads to a low rank tensor. We apply low rank optimization and solve the data completion problem using the alternating direction method of multipliers (ADMM). Computation of the tensor singular value decomposition (tSVD) is a significant component of this procedure. We improve the runtime of the tSVD algorithm by taking advantage of the even-symmetry of the Fourier transform. We demonstrate that the tensor nuclear norm (TNN) of a sampling operator is correlated with the TNN of the corresponding observation, thus producing a useful tool for determining if low rank optimization is appropriate for a given observation. Of the possible six orientations for order-3 tensors, the orientation with the most square frontal slice is best for low-rank recovery. We present our data completion results on two real datasets, the Viking Graben (offshore Norway) and Northwest Shelf 3D marine streamer data (offshore Australia).

INTRODUCTION

The data completion problem refers to recovering missing information from partial observations. It arises in many applications including for example, seismic data completion (Kreimer et al. (2013), Gao et al. (2015)), computerized tomography (Semerci et al. (2014), Mohd Sagheer and George (2019)), and video completion (Zhang et al. (2014), Long et al. (2019)). Specific situations in which data completion is implemented for exploration seismology include the need to fill in missing near and far offsets, interpolation of missing crossline data, and recovering traces when a receiver fails.

Seismic data can be naturally recorded in higher (more than two) dimensions. By arranging the data into a higher order tensor, we can use efficient and more reliable methods than by treating the data as a series of independent matrices.

There exist a variety of data completion methods including completion methods for tensors such as higher-order singular value decomposition (Kreimer and Sacchi (2012), Gao and Sacchi (2018)), parallel matrix factorization (Gao et al. (2015), Sacchi and Cheng (2017)), randomized QR decomposition (Carozzi and Sacchi (2017), Cheng and Sacchi (2015)), minimum weighted norm interpolation (Sacchi et al. (2017)), and adaptive weighted tensor completion (Ng et al. (2017)). We focus here on low rank optimization techniques, which assume the underlying true data has low rank. The rank of a matrix refers to the number of non-zero singular values in its singular value decomposition (SVD). The low rank assumption is suitable for seismic data recovery, as seismic data often contains redundancies, indicating an underlying low rank structure. It should be noted that minimizing rank directly is an NP-hard problem (Semerci et al. (2014)). Thus, the low rank requirement is often relaxed and approximated by the summation of singular values (called the nuclear norm).

In particular, we examine data completion techniques for 3D data organized as a tensor. Key concepts for matrix completion have generalized notions for higher order tensors. The tensor singular value decomposition (tSVD) generalizes the matrix SVD (Kilmer and Martin (2011)). It is used to decompose a tensor into a product of orthogonal tensors and a diagonal tensor, containing the singular values. The sum of a tensor's singular values is the tensor nuclear norm (TNN), which we use as our relaxation for the concept of tensor rank.

This paper contributes three new results for data completion of real-valued 3D data. We first present a cost improved tSVD algorithm, which makes use of the even-symmetry of the Fourier transform. The inner loop of the tSVD algorithm computes the matrix SVD for each frontal slice. A frontal slice is formed by slicing a tensor along its first and second dimensions (see Figure 1). For an $n_1 \times n_2 \times n_3$ tensor, we compute the conjugate for $\mathcal{O}(\lfloor \frac{n_3-1}{2} \rfloor)$ many slices thereby, reducing the SVD computations from $\mathcal{O}(n_3)$ to $\mathcal{O}(n_3 - \lfloor \frac{n_3-1}{2} \rfloor)$.

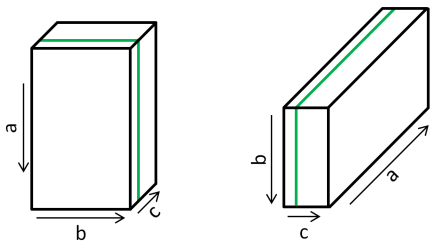

Figure 1: Visualization of two orientations for the same data. The green lines on the left highlight a frontal slice. On the right the green lines highlight a lateral slice.

The structure of the sampling operator (e.g. random or regular) has previously been shown to impact recovery (Kumar et al. (2015), Zhang and Lumley (2019), Ng et al. (2017), Popa et al. (2019)). Low-rank optimization performs better when the observation has higher rank than the ground truth. Our second contribution in this paper is demonstrating that the TNN of the sampling operator is correlated with the TNN of its observation, i.e. sampling operators with large norm tend to produce observations of large norm. This correlation is useful for predicting the success of applying low rank optimization given only the sampling operator.

Third, we present novel insight in understanding which orientation leads to the best low-rank recovery. Previous research discussed in general the dependence of algorithms, and hence recovery, on the tensor's orientation (Martin et al. (2013)). We clarify here the optimal tensor orientation for low rank

recovery specifically. For 3D data, there are six tensor orientations, Figure 1 illustrates two orientations. A rectangular matrix with dimensions $n_1 \times n_2$ has $\min(n_1, n_2)$ many singular values. Hence, the rank of the matrix is bounded by its smallest dimension. The orientation exhibiting the most square frontal slices, or 'most square' orientation, allows for observations to have higher tensor rank, resulting in better recovery. Furthermore, the 'most square' orientation coincides with the orientation of minimal runtime (Popa et al. (2019)).

TSVD SPEED-UP USING EVEN SYMMETRY

For real-valued order-3 tensors, the runtime of the tSVD algorithm can be reduced by making use of the even-symmetric property of the Fourier transform. For a vector $x = [x_1\ x_2\ \dots\ x_N]$, let $\hat{x} = \text{fft}(x)$ be the result of the fast Fourier transform (fft) applied to x. If x is real-valued then

$$\hat{x}_k = \hat{x}^*_{[-(k-1)\ \bmod N]+1},\ \forall k \in \{1,2,\dots,N\}. \tag{1}$$

This relation means that $\lfloor \frac{N-1}{2} \rfloor$ of the values in the transformed vector $\hat{x}$ are complex conjugates of the other values in the transformed vector. The tSVD algorithm first computes the fft of each tube (a vector in the third dimension). The resulting transformed tensor has even-symmetry along its frontal slices, which can be seen in Figure 2.

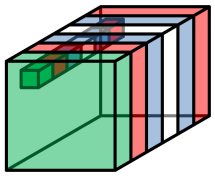

Figure 2: Visualization of an even symmetric tube in an $n_1 \times n_2 \times 6$ tensor. Slices of the same color are entry-wise complex conjugates.

Frontal slice k is the complex conjugate (entry-wise) to slice $[-(k-1)\ \bmod N]+1$. These slices have the same set of singular values in their SVD. For an $n_1 \times n_2 \times n_3$ tensor, we can compute the matrix SVD for the first $\lfloor \frac{n_3}{2} \rfloor$ slices and store the conjugate result in each corresponding symmetric slice. It was shown in Popa et al. (2019) that computing the SVD for all n_3 slices costs $\mathcal{O}((n_1^2+n_2^2)n_2n_3)$, whereas calculating the SVD and conjugate for the first $\lfloor \frac{n_3}{2} \rfloor$ slices costs $\mathcal{O}((n_1^2+n_2^2)[n_2n_3-(n_2-1)\lfloor \frac{n_3-1}{2} \rfloor])$.

Algorithm 1 outlines the accelerated tSVD, with the improvement highlighted in red. In our pseudocode, we use the notation $\text{fft}(\mathbf{X}, [], 3)$ to denote the Fourier transform in the tubal direction, ifft to denote the inverse Fourier transform, $\mathbf{X}(:,:,i)$ to denote the ith frontal slice of $\mathbf{X}$, and $\bar{\mathbf{X}}$ to denote the entry-wise conjugate of $\mathbf{X}$. In Figure 3 we compare the empirical runtime of the original and improved tSVD algorithms applied to two datasets: the Viking Graben, offshore Norway, and the Northwest Shelf 3D marine streamer data, obtained from the west coast of Australia. For the Viking Graben data we observe a 48% reduction in runtime of the inner loop and for the Northwest Shelf we observe a 35% improvement.

Algorithm 1: Efficient tSVD for real-valued order-3 tensors

Input: $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$;
$\hat{\mathbf{X}} \leftarrow \text{fft}(\mathbf{X}, [], 3)$;
$[\hat{\mathbf{U}}(:,:,1), \hat{\mathbf{S}}(:,:,1), \hat{\mathbf{V}}(:,:,1)] = \text{svd}(\hat{\mathbf{X}}(:,:,1))$;
for *i = 1 to* $\lfloor \frac{n_3}{2} \rfloor$ **do**
  $[\hat{\mathbf{U}}(:,:,i+1), \hat{\mathbf{S}}(:,:,i+1), \hat{\mathbf{V}}(:,:,i+1)] = \text{svd}(\hat{\mathbf{X}}(:,:,i+1))$;
  $j = n_3 - i + 1$;
  $\hat{\mathbf{S}}(:,:,j) = \hat{\mathbf{S}}(:,:,i+1)$;
  $\hat{\mathbf{U}}(:,:,j) = \bar{\hat{\mathbf{U}}}(:,:,i+1)$;
  $\hat{\mathbf{V}}(:,:,j) = \bar{\hat{\mathbf{V}}}(:,:,i+1)$;
end
$\mathbf{U} \leftarrow \text{ifft}(\hat{\mathbf{U}}, [], 3)$;
$\mathbf{S} \leftarrow \text{ifft}(\hat{\mathbf{S}}, [], 3)$;
$\mathbf{V} \leftarrow \text{ifft}(\hat{\mathbf{V}}, [], 3)$;

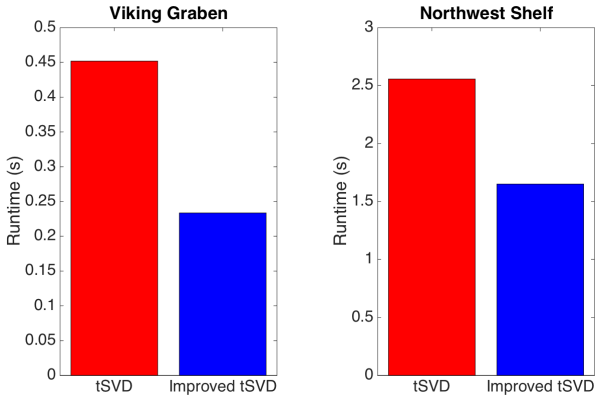

Figure 3: Comparison of empirical runtime of the inner loop of the original and improved tSVD applied to the Viking Graben (left) and the Northwest Shelf datasets (right).

SAMPLING OPERATOR RESULTS

The data completion problem can be expressed mathematically as,

$$\mathbf{Y} = A(\mathbf{X}). \tag{2}$$

From (2) we seek to reconstruct the true data $\mathbf{X}$, provided with observation $\mathbf{Y}$ and sampling operator A. We consider the problem without denoising for simplicity.

We can treat the sampling operator as a tensor containing zeroes and ones; a zero where data is missing and a one where data is sampled. The sampling operator applied to a data tensor requires taking the Hadamard product (entrywise multiplication). In our experiments, we remove 60% of the data using our sampling operator. We compare three types of sampling operators: regular, random, and random column. *Regular sampling* decimates traces with regular spacing. *Random sampling* selects data at random to be removed. *Random columns* removes traces (columns) selected at random. We note that fully random sampling does not occur in real world applications, but

is presented for mathematical completeness. We apply each operator to our 'ground truth' data to produce an observation.

For a tensor we compute the tensor nuclear norm (TNN) as,

$$||\mathbf{X}||_{\text{TNN}} = \sum_{i=1}^{N} ||\hat{\mathbf{X}}(:,:,i)||_{\text{nuc}}, \tag{3}$$

where $\hat{\mathbf{X}}$ is the Fourier Transform of $\mathbf{X}$ and $||\cdot||_{nuc}$ is the matrix nuclear norm, calculated by summing the singular values.

We compare the TNN of each operator and its corresponding observation in Figure 4(a) for the Viking Graben data, observing that operators with larger norm produce observations of larger norm. Both random sampling operators have greater TNN values than the identity operator, and produce observations of larger TNN than the ground truth. Conversely, the regular operator has smaller TNN than the identity and produces an observation with TNN value less than the true data's. Thus comparing the TNN of a sampling operator to the TNN of the identity operator can indicate if the observation has higher rank than the ground truth, indicating whether low rank optimization is suitable for any given observation.

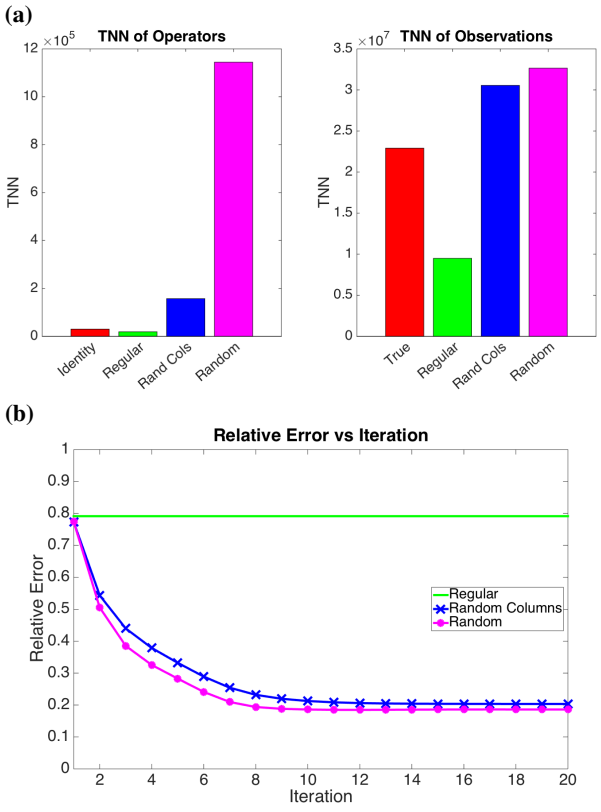

Figure 4: **(a)** TNN of sampling operators (left) and observations (right) for the Viking Graben data. **(b)** Error for each sampling operator applied to the Viking Graben data across 20 iterations of ADMM.

To demonstrate the effect of the sampling operator on the recovered result we solve the data completion problem. We seek to recover $\mathbf{X}$ by minimizing TNN (our relaxation of tensor rank), which results in the optimization problem,

$$\min ||\mathbf{X}||_{\text{TNN}}, \quad \text{s.t.}\, \mathbf{Y} = A(\mathbf{X}). \tag{4}$$

To solve, we apply the alternating direction method of multipliers (ADMM) (Boyd et al. (2010), Ely et al. (2015)). ADMM splits the TNN penalty and data matching constraint, $\mathbf{Y} = A(\mathbf{X})$, into two subproblems, each having a closed-form solution.

In Figure 4(b) we plot the error in recovery for each sampling scheme and in Figure 5(b) we compare the recovered results. Since the observation produced by regular sampling has lower TNN than the underlying data, low rank optimization is unable to recover the data, and the algorithm produces an error around 80%. In contrast, the two random sampling schemes converge with error near 20%.

ORIENTATION RESULTS

The tSVD algorithm is dependent on orientation, resulting in different sets of singular values. 3D tensors have six orientations. Orientations for which the first two dimensions are transposed have the same singular values, meaning we can consider only three distinct orientations, $a \times b \times c$, $a \times c \times b$ and $b \times c \times a$. The set of singular values determines the TNN and ultimately, impacts the recovered result.

It has been shown that the tSVD runtime for an $n_1 \times n_2 \times n_3$ tensor is $\mathcal{O}(n_3[(n_1+n_2)^2 \log(n_3) + (n_1^2+n_2^2)n_2])$, the sum of costs for the ftt, SVD, and ifft steps (Popa et al. (2019)). Choosing n_2 as the smallest dimension and n_3 as the largest minimizes the runtime, i.e. $n_2 \leq n_1 \leq n_3$. Additionally, if n_1 is relatively closer to n_2 than to n_3, then this orientation is the 'most square.'

To demonstrate which orientation gives the best recovery, we apply the data completion algorithm to CMP gathers from the Northwest Shelf dataset. Each gather has 1732 time steps and 182 receivers. Using 20 gathers, we permute the 3D data tensor for the three distinct orientations and compare recovery after 20 iterations of ADMM. Figure 6(a) contrasts the error for each orientation. Bars of the same color have their first and second dimensions transposed, and produce the same result. The 'most square' orientations (leftmost in Figure 6(a)) have a relative error around 13.5%, an 11% improvement over the other orientations.

To better see the effect of orientation on the true data, we compare the normalized cumulative sum of singular values for the three distinct orientations in Figure 6(b). Low rank tensors can be identified if a small percent of singular values contain the majority of the information in the data. The 'most square' orientation exhibits the most low rank behavior since for any given percent of singular values, the sum is a greater percent of the cumulative total than for the other orientations.

While orientation plays an important role in recovery, the sampling operator has a more dominant effect. We compare the final recovery error for each orientation and each sampling operator in Table 1. Regular sampling fails for all three orientations. For the two random sampling schemes we note that the most square orientation performs best.

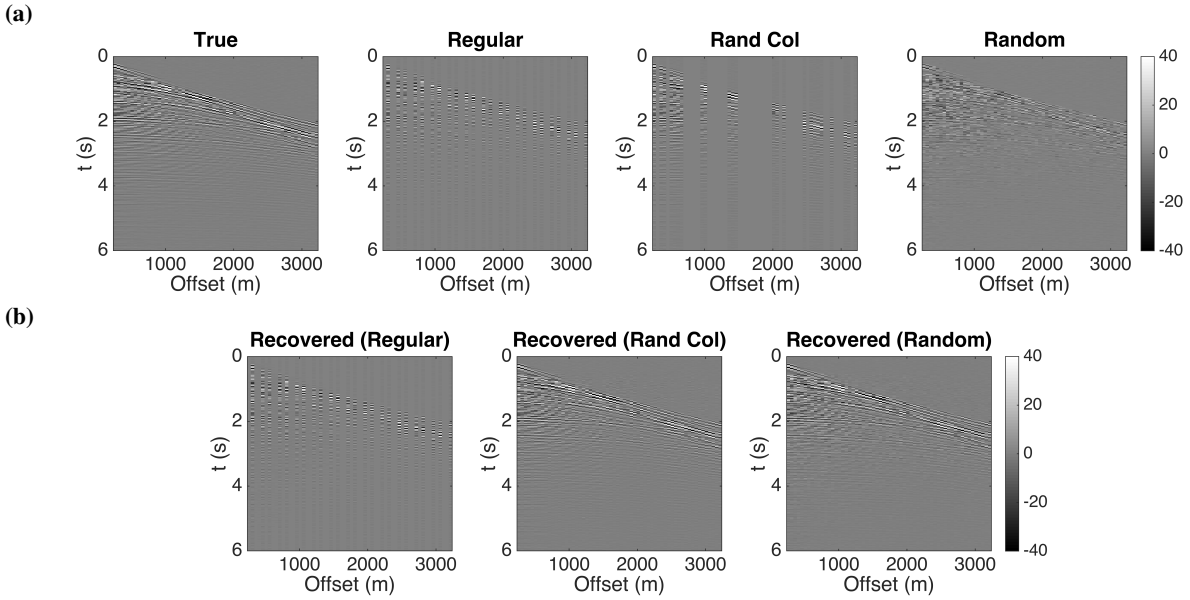

Figure 5: **(a)** (left to right) True data and observations from the Viking Graben data for regular, random column, and fully random sampling. **(b)** Recovered result for each operator.

Northwest Shelf Recovery Error			
Sampling Operator	$1732 \times 182 \times 20$	$1732 \times 20 \times 182$	$182 \times 20 \times 1732$
Random	.27867	.24913	.13565
Rand Col	.37463	.30951	.27634
Regular	.74862	.61836	.69592

Table 1: Error in recovery of the Northwest Shelf data for each sampling operator and each orientation. The third column corresponds with the 'most square' orientation.

CONCLUSIONS

For real-valued order-3 tensors, the tSVD algorithm runtime can be reduced by utilizing the even-symmetric property of the Fourier transform. The norm of the sampling operator indicates the effect sampling will have on the observation, which in turn affects recovery. Our orientation results support that the 'most square' orientation produces the best low rank recovery, while also minimizing runtime.

ACKNOWLEDGMENTS

This research is partially supported by the sponsors of the UT Dallas "3D+4D Seismic FWI" research consortium and by NSF CAREER Award #1846690. We thank BHP for providing the Northwest Shelf 3D+4D marine streamer data. We thank Mobil Oil (now ExxonMobil) for making publicly available the seismic data set from the Viking Graben. We also thank Wei Zhou and David Lumley for providing the data used in our experiments.

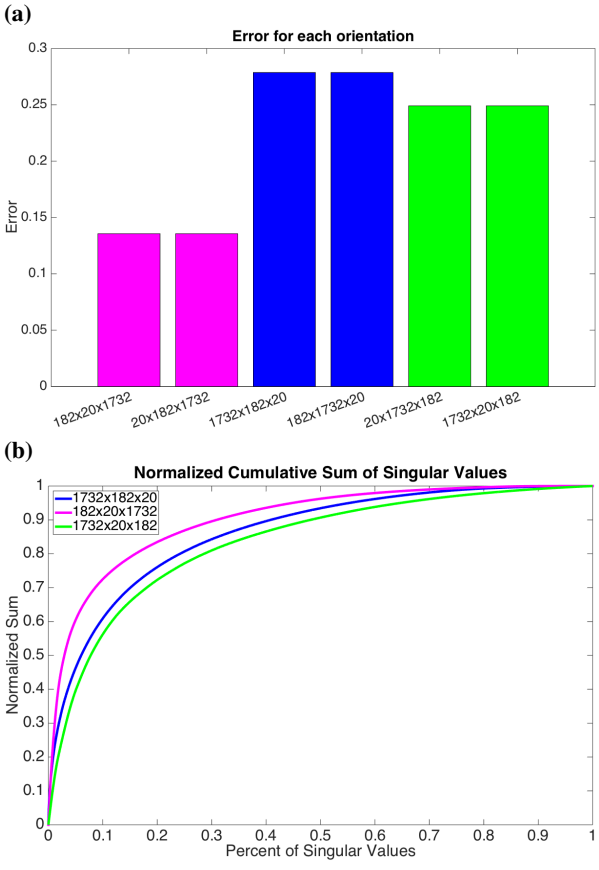

Figure 6: **(a)** Error for each orientation of the Northwest Shelf data after 20 iterations of ADMM. **(b)** The normalized cumulative sum of the singular values for the three distinct orientations of the Northwest Shelf data.

REFERENCES

Boyd, S., N. Parikh, E. Chu, B. Peleato, and J. Eckstein, 2010, Distributed optimization and statistical learning via the alternating direction method of multipliers: Foundations and Trends in Machine Learning, **3**, 1–122, doi: https://doi.org/10.1561/2200000016.

Carozzi, F., and M. D. Sacchi, 2017, 5D seismic reconstruction via PMF with randomized QR decomposition: 87th Annual International Meeting, SEG, Expanded Abstracts, 4251–4256, doi: https://doi.org/10.1190/segam2017-17782228.1.

Cheng, J., and M. D. Sacchi, 2015, A fast rank-reduction algorithm for 3D deblending via randomized QR decomposition: 85th Annual International Meeting, SEG, Expanded Abstracts, 3830–3835, doi: https://doi.org/10.1190/segam2015-5850767.1.

Ely, G., S. Aeron, N. Hao, and M. E. Kilmer, 2015, 5D seismic data completion and denoising using a novel class of tensor decompositions: Geophysics, **80**, V83–V95, doi: https://doi.org/10.1190/geo2014-0467.1.

Gao, J., A. Stanton, and M. D. Sacchi, 2015, Parallel matrix factorization algorithm and its application to 5D seismic reconstruction and denoising: Geophysics, **80**, 173–187, doi: https://doi.org/10.1190/geo2014-0594.1.

Gao, W., and M. D. Sacchi, 2018, PP-wave and PS-wave 5D reconstruction and denosing to assist registration: 88th Annual International Meeting, SEG, Expanded Abstracts, 2397–2401, doi: https://doi.org/10.1190/segam2018-2997862.1.

Kilmer, M. E., and C. D. Martin, 2011, Factorization strategies for third-order tensors: Linear Algebra and its Applications, **435**, 641–658, doi: https://doi.org/10.1016/j.laa.2010.09.020.

Kreimer, N., and M. D. Sacchi, 2012, A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation: Geophysics, **77**, 113–122, doi: https://doi.org/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**, 273–284, doi: https://doi.org/10.1190/geo2013-0022.1.

Kumar, R., C. Da Silva, O. Akalin, A. Y. Aravkin, H. Mansour, B. Recht, and F. J. Herrmann, 2015, Efficient matrix completion for seismic data reconstruction: Geophysics, **80**, V97–V113, doi: https://doi.org/10.1190/geo2014-0369.1.

Long, Z., Y. Liu, L. Chen, and C. Zhu, 2019, Low rank tensor completion for multiway visual data: Signal Processing, 301–316, doi: https://doi.org/10.1016/j.sigpro.2018.09.039.

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: https://doi.org/10.1137/110841229.

Mohd Sagheer, S. V., and S. N. George, 2019, Denoising of low-dose CT images via low-rank tensor modeling and total variation regularization: Artificial Intelligence in Medicine, **94**, 1–17, doi: https://doi.org/10.1016/j.artmed.2018.12.006.

Ng, M. K.-P., Q. Yuan, L. Yan, and J. Sun, 2017, An adaptive weighted tensor completion method for the recovery of remote sensing images with missing data: IEEE Transactions on Geoscience and Remote Sensing, **55**, 3367–3381, doi: https://doi.org/10.1109/TGRS.2017.2670021.

Popa, J., S. Minkoff, and Y. Lou, 2019, Improving seismic data completion and efficiency using tensors: 89th Annual International Meeting, SEG, Expanded Abstracts, 4034–4038, doi: https://doi.org/10.1190/segam2019-3215431.1.

Sacchi, M. D., and J. Cheng, 2017, 5D reconstruction via robust tensor completion: GeoConvention, 1–5, doi: https://doi.org/10.1190/geo2018-0109.1.

Sacchi, M. D., G. Garabito, H. Schots, and J. Caldeira, 2017, Preconditioning and denoising prestack onshore seismic data via 5D reconstruction: Application to 3D data from the Parnaiba Basin: 15th International Congress of the Brazilian Geophysical Society & EXPOGEF, 1384–1388.

Semerci, O., N. Hao, M. E. Kilmer, and E. L. Miller, 2014, Tensor-based formulation and nuclear norm regularization for multienergy computed tomography: IEEE Transactions on Image Processing, **23**, 1678–1693, doi: https://doi.org/10.1109/TIP.2014.2305840.

Zhang, M., and D. Lumley, 2019, Reconstruction of 3D seismic data from sparse random OBN acquisition by compressive sensing: 89th Annual International Meeting, SEG, Expanded Abstracts, 127–130, doi: https://doi.org/10.1190/segam2019-3213090.1.

Zhang, Z., G. Ely, S. Aeron, N. Hao, and M. Kilmer, 2014, Novel methods for multilinear data completion and denoising based on tensor-SVD: IEEE Conference on Computer Vision and Pattern Recognition, 3842–3849, doi: https://doi.org/10.1109/CVPR.2014.485.