

Efficient Tensor Completion Methods for 5-D Seismic Data Reconstruction: Low-Rank Tensor Train and Tensor Ring

Dawei Liu^{ID}, Mauricio D. Sacchi^{ID}, Member, IEEE, and Wenchao Chen^{ID}

Abstract—Five-dimensional seismic reconstruction is receiving increasing attention and can be viewed as a tensor completion problem, which involves reconstructing a low-rank tensor from a partially observed tensor. Tensor train (TT) decomposition and tensor ring (TR) decomposition are two powerful tensor networks for solving this problem. However, updating core tensors leads to high computational costs in practical applications. We propose two efficient methods to exploit low TT rank and low TR rank structures by theoretically establishing the relationship between tensor ranks and matrix unfoldings, respectively. Specifically, the former uses a well-balanced matricization scheme, and the latter uses a tensor circular unfolding. Furthermore, we use the randomized parallel matrix factorization (PMF) to accelerate the solution of these problems. Both synthetic and real data experiments demonstrate that the proposed algorithm can also achieve remarkable reconstruction performance; in the meantime, the computational cost is significantly reduced.

Index Terms—Low-rank, matrix factorization (MF), seismic data reconstruction, tensor completion, tensor ring (TR), tensor train (TT).

I. INTRODUCTION

DUCE to certain obstacles in the seismic acquisition process, such as irregular surface conditions and equipment limitations, seismic data are often sampled irregularly along spatial coordinates. The missing traces in seismic records negatively affect subsequent processing, imaging, and interpretation involving multichannel deconvolution [1], velocity analysis [2], full-waveform inversion [3], simultaneous-source separation [4], and fault detection [5]. Consequently, seismic data reconstruction is an ongoing and vital problem, which has attracted considerable attention from academia and industry in the past several decades.

A number of techniques have been developed to reconstruct seismic data. Prediction filters are the first category to achieve effective reconstruction results [6]–[8]. By taking advantage of the predictability of linear events in the

Manuscript received January 27, 2022; revised April 22, 2022; accepted May 24, 2022. Date of publication May 30, 2022; date of current version June 14, 2022. This work was supported in part by the National Key Research and Development Program of China under Grant 2021YFA0716904 and in part by the National Natural Science Foundation of China under Grant 41974131 and Grant 41774135. The work of Dawei Liu was supported by the China Scholarship Council. (*Corresponding author: Dawei Liu.*)

Dawei Liu and Wenchao Chen are with the School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: liudawei2015@stu.xjtu.edu.cn; wencchen@xjtu.edu.cn).

Mauricio D. Sacchi is with the Department of Physics, University of Alberta, Edmonton, AB T6G 2E1, Canada (e-mail: msacchi@ualberta.ca).

Digital Object Identifier 10.1109/TGRS.2022.3179275

frequency-space domain, aliased high-frequency data can be interpolated by filters derived from nonaliased low-frequency data. Therefore, antialiasing reconstruction is a significant benefit of these methods. Unfortunately, the requirement that seismic observations should be equally sampled limits the practical applicability of this technique. Moreover, they operate in small windows where one can assume that data are composed of a few dips. The second category comprises transform-based methods, which play a pivotal role in addressing the issue of recovering missing traces. The Fourier transform with simplicity constraints, such as band-limited wavenumber-domain regularization [9]–[11] and a sparsity-promoting regularization of spectral amplitudes [12]–[16], is widely used by industry. This transform is particularly useful in 5-D seismic volume reconstruction [17]–[19]. The Radon transform [20]–[22], the wavelet transform [23], the shearlet transform [24], and the curvelet transform [25]–[27] are also used to interpolate unknown traces. Data-driven techniques are the third category of reconstruction methods, which aims to better fit the data and preserve the subtle features. Dictionary learning methods, involving the K-singular value decomposition [28], [29] and the data-driven tight frame method [30], can learn the features directly from the input data in a data-driven manner, thus becoming more adaptive than the aforementioned fixed-basis transform. Nevertheless, the high computational complexity and extreme parameter sensitivity render it unaffordable when dealing with high-dimensional problems. Recent advances in deep learning make it a powerful tool to extract features through data-driven training [31], [32], which is therefore introduced to seismic data interpolation. Although they have achieved promising results, the following issues still need to be further addressed. Supervised deep learning [33]–[35] acquires the prior knowledge by learning the training dataset, which is a thorny issue for real data. Unsupervised deep learning [36]–[38] can avoid the preparation of training datasets, but it needs much more testing time, especially for large-scale datasets since the network usually needs to be retrained for each input. 5-D pre-stack data are highly complicated, making unsupervised learning more challenging. In addition, even though these methods can easily access the GPU, they still cannot burden the computational costs beyond 3-D applications. In other words, they are inadequate to leverage the comprehensive high-dimensional relationships inherent in 5-D seismic data.

An alternative to constraining the reconstruction problem is the assumption that complete data are low-rank and an increase in missing traces and noise results in a higher rank [39], [40]. There are two basic approaches currently being adopted in research into rank-reduction-based methods: matrix-based and tensor-based approaches. The former first rearranges multidimensional seismic temporal-frequency slices into multilevel block Hankel/Toeplitz matrices or texture matrices and then applies a rank reduction algorithm to recover the data [41]–[44]. The latter regards multichannel seismic data as multilinear arrays or tensors and directly operates on the tensor via dimensionality reduction techniques, such as the high-order SVD [45], Tucker decomposition [46], and tensor SVD [47]. Gao *et al.* [48] adopt the parallel matrix factorization (PMF) algorithm [49] to reconstruct a low-rank tensor structure from undersampled seismic data. This technique performs matrix factorizations to different tensor unfoldings, thus dramatically decreasing the computational complexity by avoiding the computation of SVD. Furthermore, parallel square matrix factorization (PSMF) is proposed by adopting an elaborate unfolding procedure to enhance the reconstruction performance and quality [50]. These methods preserve the seismic waveforms better since all physical dimensions of seismic data are taken into consideration, demonstrating the inherent potential of tensor algebra for the development of new reconstruction techniques.

More recently, tensor-network-based tensor decompositions have gained growing popularity and shown remarkable promise in dealing with higher order, especially beyond third-order, tensors [51]–[53]. Successful applications include hyperspectral image super-resolution [54] and denoising [55], especially tensor completion [56], [57]. As one of the most representative among them, tensor train (TT) decomposition [58] has been intensively investigated recently in numerous tasks due to its high data compression ability and computational efficiency [59]–[61]. As a generalization of TT decomposition, tensor ring (TR) decomposition breaks down a tensor as several cyclically contracted third-order tensors [62]. Besides the dominant properties of TT, TR has enhanced compressibility and flexible cores that are circularly shiftable, thus showing promising potential for seismic reconstruction. In general, SVD [58], alternating least-squares method [63], and the gradient descent algorithm [64] are used to identify latent core tensors of TT and TR decompositions for exploiting the low-rank structure. However, despite the existence of numerous optimization algorithms, explicit TT and TR decompositions are still intractable for large-scale applications due to the high computational cost. Note that the above assumptions underlying tensor completion are that TT and TR core tensors are low-rank, regardless of whether they are explicitly expressed. Therefore, we exploit the low-rank structure in an implicit way by minimizing the TT rank or TR rank. In other words, parallel low-rank matrix factorization is performed to different tensor unfoldings, substantially reducing the computational cost in estimating missing entries.

In this article, we propose two efficient tensor completion methods based on TT rank and TR rank minimization. These methods lead to a new type of algorithm for seismic data

reconstruction. TT rank and TR rank have a remarkable capacity for characterizing the correlations between different modes in higher order tensors. To achieve it, TT rank and TR rank each applies a new balanced unfolding scheme to the tensor, respectively. Then, with PMF, low TT rank completion and low TR rank completion are accomplished in an efficient manner. Meanwhile, a randomized algorithm, *sketching*, is further applied to accelerate the proposed method. Both synthetic and field experiments are conducted to demonstrate that the proposed algorithm can significantly accelerate the 5-D seismic reconstruction process and achieve outstanding results. The main contributions of this work can be summarized as follows.

- 1) We propose a low-rank 5-D seismic data reconstruction method based on PMF via TT (PMF-TT) rank minimization. To the best of our knowledge, this is the first time TT is applied to seismic data processing tasks.
- 2) As the TR rank is an extension of the TT rank, PMF via TR (PMF-TR) rank minimization is also studied, associating with a well-balanced unfolding scheme. Furthermore, due to the symmetry of circular unfolding, we find that only half of the PMF-TR unfoldings can produce similar results. For distinguishing it from the original PMF-TR, this new version is named PMF-TRH, which requires lower computational resources than PMF-TR.
- 3) A randomized strategy for accelerating the aforementioned methods is introduced, which can significantly reduce the computation time. As far as we know, the study presented in this article is the first investigation to accelerate PMF-based reconstruction.

The remainder of the article is organized as follows. In Section II, we first describe the tensor basics and notations used throughout this article. Then, the connections of TT/TR ranks between the tensor decomposition and the tensor unfoldings are introduced. Section III proposes two categories of 5-D seismic data reconstruction methods, and a randomized algorithm is used to accelerate the proposed methods. The results of the reconstruction experiments on synthetic and field seismic data are presented in Section IV, verifying that the proposed method offers superior performance and is computationally efficient. Finally, the conclusion is presented in Section V.

II. NOTATIONS AND PRELIMINARIES

A. Notations

In this article, we adopt the notations described in [65]. Scalars (e.g., $x, X \in \mathbb{R}$) are denoted by standard letters. A vector, also known as a first-order tensor, is inscribed in a boldface lowercase letter, e.g., $\mathbf{x} \in \mathbb{R}^I$. Boldface capital letters denote matrices, e.g., $\mathbf{X} \in \mathbb{R}^{I \times J}$. Tensors of order $N \geq 3$ with size $I_1 \times I_2 \times \dots \times I_N$ are indicated by calligraphic letters, e.g., $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$. An element of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ is denoted by $\mathcal{X}(i_1, i_2, \dots, i_N)$ or $x_{i_1 i_2 \dots i_N}$, where (i_1, i_2, \dots, i_N) is the index. The inner product of two tensors \mathcal{X}, \mathcal{Y} with the same size $\mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ is defined as $\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{i_1} \sum_{i_2} \dots \sum_{i_N} x_{i_1 i_2 \dots i_N} y_{i_1 i_2 \dots i_N}$. Furthermore, the Frobenius norm of \mathcal{X} is defined by $\|\mathcal{X}\|_F = ((\mathcal{X}, \mathcal{X}))^{1/2}$.

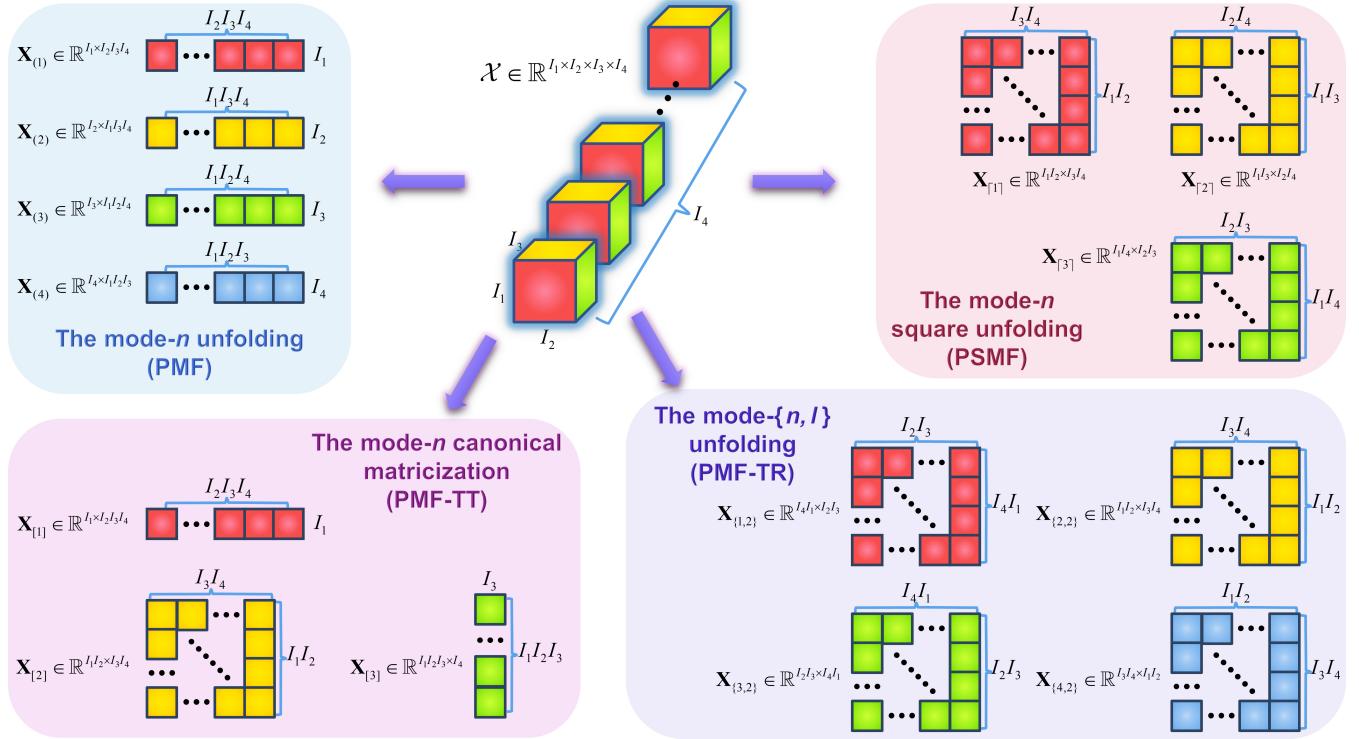


Fig. 1. Illustration of different tensor unfolding schemes using a fourth-order tensor \mathcal{X} . The mode- n canonical matricization is used for PMF via TT rank minimization. The mode- $\{n, l\}$ unfolding is adopted in PSMF via TR rank minimization.

Assembling the entries of a tensor into a matrix is called the tensor unfolding or matricization, and there are several varieties as shown in Fig. 1. The mode- n unfolding of tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is a common practice, denoted by $\mathbf{X}_{(n)} \in \mathbb{R}^{I_n \times I_1 \cdots I_{n-1} I_{n+1} \cdots I_N}$. An alternative mode- n unfolding of tensor \mathcal{X} [62] is denoted by $\mathbf{X}_{<n>} \in \mathbb{R}^{I_n \times I_{n+1} \cdots I_N I_1 \cdots I_{n-1}}$, which is often used in TR operations. The mode- n canonical matricization of \mathcal{X} is also frequently used, defined by $\mathbf{X}_{[n]} \in \mathbb{R}^{(\prod_{j=1}^n I_j) \times (\prod_{j=n+1}^N I_j)}$. The mode- $\{n, l\}$ unfolding of \mathcal{X} , which is a circular unfolding, is characterized by $\mathbf{X}_{\{n, l\}} \in \mathbb{R}^{(\prod_{j=n-l+1}^n I_j) \times (\prod_{j=n+l+1}^{n-l} I_j)}$. In addition, unfolding \mathcal{X} into almost square matrices is used in PSMF, which is expressed by $\mathbf{X}_{[n]} = \text{reshape}_n[\mathcal{X}]$. Matrix folding, or tensorization, is the inverse operation of unfolding in which matrices are converted into higher order tensors. Take mode- n unfolding as an example, the corresponding mode- n folding is denoted by $\text{fold}_{(n)}$ such that $\text{fold}_{(n)}(\mathbf{X}_{(n)}) = \mathcal{X}$. For any pair of folding and unfolding, we have $\text{fold}_n(\text{unfold}_n(\mathcal{X})) = \mathcal{X}$.

B. Tensor Train Rank

A previous study of PMF [48] has investigated the relationships between the Tucker rank and the mode- n tensor unfolding, i.e., $\text{rank}_{\text{Tucker}}(\mathcal{X}) = [\text{rank}(\mathbf{X}_{(1)}), \dots, \text{rank}(\mathbf{X}_{(N)})]$, and the Tucker rank minimization approach has been successfully applied to tensor completion. As parameters in TT format presentation is much smaller than that in Tucker representations [66], minimizing the TT rank should allow us to handle more complicated tensor completion situations. Here, we introduce the TT rank which depends on TT decomposition and is a generalization of the Tucker rank.

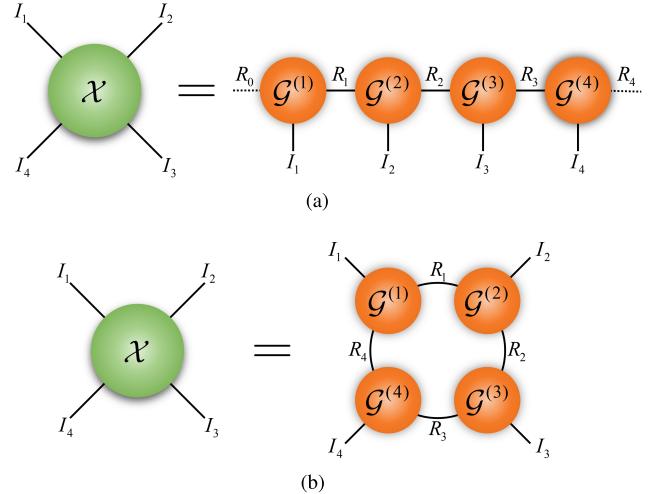


Fig. 2. Illustration of tensor decomposition with tensor diagrams. (a) TT decomposition, where $R_0 = R_N = 1$. (b) TR decomposition.

TT decomposition is to decompose a tensor into a series of tensor cores, as shown in Fig. 2(a). Each tensor core is a three-way tensor. Specifically, the minimal TT decomposition of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is stated in the following form:

$$\mathcal{X} = \ll \mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \dots, \mathcal{G}^{(N)} \gg \quad (1)$$

where $\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \dots, \mathcal{G}^{(N)}$ refers to a sequence of three-way tensor cores of size $1 \times I_1 \times R_1, R_1 \times I_2 \times R_2, \dots, R_{N-1} \times I_N \times 1$, respectively. $\ll \cdot \gg$ denotes multilinear products of the above cores. The sequence $\{R_1, R_2, \dots, R_{N-1}, 1\}$, denoted as $\text{rank}_{\text{TT}}(\mathcal{X})$, is the so-called TT rank which imposes limits on

the size of each TT core. The elementwise relationship of TT decomposition and tensor \mathcal{X} can be written as the following index form:

$$\mathcal{X}(i_1, i_2, \dots, i_N) = \mathbf{G}_{i_1}^{(1)} \times \mathbf{G}_{i_2}^{(2)} \times \dots \times \mathbf{G}_{i_N}^{(N)} \quad (2)$$

where $\mathbf{G}_{i_n}^{(n)} = \mathcal{G}^{(n)}(:, i_n, :) \in \mathbb{R}^{R_{n-1} \times R_n}$ denotes the i_n th slice of $\mathcal{G}^{(n)}$. Naturally, we can obtain the TT rank after minimal TT decomposition. However, obtaining the TT cores of minimal TT decomposition is often computationally intensive. Therefore, finding an efficient way to characterize the TT rank is essential.

An earlier study [51] has theoretically revealed the connection between the mode- n canonical matricization of \mathcal{X} and TT cores by the following equation:

$$\mathbf{X}_{[n]} = \mathbf{G}_{[n]}^{\leq n} \mathbf{G}_{(1)}^{>n} \quad (3)$$

where $\mathcal{G}^{\leq n} \in \mathbb{R}^{I_1 \times \dots \times I_n \times R_n}$ and $\mathcal{G}^{>n} \in \mathbb{R}^{R_n \times I_{n+1} \times \dots \times I_N}$ are the partial contracted products. Accordingly, the rank of the mode- n canonical matricization and TT rank satisfies

$$\begin{aligned} \text{Rank}(\mathbf{X}_{[n]}) &\leq \min \left\{ \text{Rank}(\mathbf{G}_{[n]}^{\leq n}), \text{Rank}(\mathbf{G}_{(1)}^{>n}) \right\} \\ &\leq R_n. \end{aligned} \quad (4)$$

More specifically, the rank of $\mathbf{X}_{[n]}$ is the lower bound of R_n . In addition, a seminal study [67] demonstrates that $\text{Rank}(\mathbf{X}_{[n]}) = R_n$ holds true for a minimal TT decomposition, where $n = 1, \dots, N - 1$. Consequently, $\text{Rank}(\mathbf{X}_{[n]})$ is frequently used to represent the TT rank when solving TT rank minimization problems.

C. Tensor Ring Rank

In the TT format, there are two matrices in the first and last positions, producing large middle cores and small border factors. This limitation leads to unbalanced unfolding and negatively impacts its representational ability and flexibility, especially for low-order tensors such as seismic data. As a more general decomposition technique, TR decomposition represents the tensor \mathcal{X} by circular multilinear products over a sequence of three-way latent core tensors $\{\mathcal{G}^{(n)}\}_{n=1}^N$, as depicted in Fig. 2(b). In contrast to TT cores, $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(N)}$ in TR cores have sizes of $R_0 \times I_1 \times R_1$ and $R_{N-1} \times I_N \times R_N$, respectively, where $R_0 = R_N$. The elements of \mathcal{X} under TR decomposition can be expressed as

$$\mathcal{X}(i_1, i_2, \dots, i_N) = \text{Tr} \left\{ \prod_{n=1}^N \mathbf{G}_{i_n}^{(n)} \right\} \quad (5)$$

where $\text{Tr}(\cdot)$ stands for the matrix trace operation. Similar to TT rank, the sequence $\{R_1, R_2, \dots, R_{N-1}, R_N\}$, denoted as $\text{rank}_{\text{TT}}(\mathcal{X})$, is TR rank.

According to the following equation proposed in [62]:

$$\mathbf{X}_{<n>} = \mathbf{G}_{(2)}^{(n)} \left(\mathbf{G}_{<2>}^{(\neq n)} \right)^T \quad (6)$$

the TR rank is less than or equal to the corresponding core tensor rank since the inequality holds for all

$n = 1, \dots, N$ as follows:

$$\begin{aligned} \text{Rank}(\mathbf{X}_{<n>}) &\leq \min \left\{ \text{Rank}(\mathbf{G}_{(2)}^{(n)}), \text{Rank}(\mathbf{G}_{<2>}^{(\neq n)}) \right\} \\ &\leq \text{Rank}(\mathbf{G}_{(2)}^{(n)}) \\ &\leq R_n, \end{aligned} \quad (7)$$

where $\mathcal{G}^{(\neq n)} \in \mathbb{R}^{R_{n+1} \times \prod_{j=1, j \neq n}^N I_j \times R_n}$ is a subchain tensor produced by the contracted products of all core tensors except the n th one. Apparently, the low-rank constraint on $\mathcal{G}^{(n)}$ can be substituted by $\text{Rank}(\mathbf{X}_{<n>})$, which is a common practice to solve TR rank minimization problems. However, this unfolding scheme still results in “long strip” matrices, which are not sufficiently balanced.

Motivated by the fact that PSMF outperforms PMF due to a balanced tensor unfolding [50], we use a circular unfolding $\mathbf{X}_{\{n,l\}}$ to represent the TR rank. Contrary to $\mathbf{X}_{<n>}$ which conducts tensor unfolding along a single mode, $\mathbf{X}_{\{n,l\}}$ unfolds \mathcal{X} along l modes $\{n-l+1, \dots, n-1, n\}$, resulting in more balanced factor matrices. Typically, l is set to be $\lfloor (N/2) \rfloor$ to generate well-balanced matrices. Similarly, this circular unfolding can maintain the low-rank property of \mathcal{X} by the inequality [68] $\text{rank}(\mathbf{X}_{\{n,l\}}) \leq R_{n-l} R_n$, establishing connections straightforwardly toward the TR rank. Therefore, $\mathbf{X}_{\{n,l\}}$ is adopted in this article to capture the low TR rank structure hidden in seismic data.

III. THEORY

This section first revisits the conventional formulation of seismic reconstruction, known as PMF. Based on PMF, we propose two types of novel approaches via TT rank and TR rank minimization optimization. Finally, a fast algorithm is used to accelerate the proposed methods.

A. Seismic Reconstruction Based on Parallel Matrix Factorization Algorithm

The midpoint-offset frequency-space domain is widely used for prestack seismic reconstruction, where the fully sampled seismic data have a low-rank structure. In this scenario, the seismic reconstruction problem can be represented as a low-rank tensor completion. Specifically, a noisy and incomplete 5-D seismic dataset is denoted by $D^{\text{obs}}(\omega, x, y, h_x, h_y)$, where x and y represent the spatial coordinates of the inline and cross-line midpoints, respectively, and h_x and h_y indicate the inline and cross-line offsets, respectively. We introduce a fourth-order tensor \mathcal{D}^{obs} as a representation of the binned seismic volume at a specific temporal frequency ω . The missing entries of \mathcal{D}^{obs} are substituted with zeros. The elementwise representation is $d_{i_1 i_2 i_3 i_4}^{\text{obs}}$, where binning indices i_1, i_2, i_3 , and i_4 correspond to the spatial coordinates x, y, h_x , and h_y , respectively. While the notation has been streamlined by dropping the dependency on ω , it is worth noting that subsequent analysis is still performed for all frequencies within the range $\omega \in [\omega_{\min}, \omega_{\max}]$.

Assume that the complete data acquired from an ideal survey are a low-rank tensor \mathcal{Z} , the mathematical relationship

that exists between \mathcal{Z} and the observed data \mathcal{D}^{obs} can be formulated as follows:

$$\mathcal{P} \circ \mathcal{Z} = \mathcal{D}^{\text{obs}} \quad (8)$$

where “ \circ ” represents the Hadamard (elementwise) product and \mathcal{P} indicates a sampling operator in which every entry complies

$$p_{i_1 i_2 i_3 i_4} = \begin{cases} 0 & \text{if } d_{i_1 i_2 i_3 i_4} \text{ is a missing entry} \\ 1 & \text{if } d_{i_1 i_2 i_3 i_4} \text{ is an observed entry.} \end{cases} \quad (9)$$

The problem of recovering the missing entries of \mathcal{D}^{obs} can be addressed via the well-known rank minimization algorithm

$$\min_{\mathcal{Z}} \text{rank}(\mathcal{Z}) \quad \text{s.t. } \mathcal{P} \circ \mathcal{Z} = \mathcal{D}^{\text{obs}}. \quad (10)$$

To recover the missing entries of \mathcal{Z} , the rank of \mathcal{Z} should be kept as low as feasible to capture the low-rank structure. Alternatively, the unconstrained form can be expressed as

$$\min_{\mathcal{Z}} \text{rank}(\mathcal{Z}) + \frac{\mu}{2} \|\mathcal{P} \circ \mathcal{Z} - \mathcal{D}^{\text{obs}}\|_F^2 \quad (11)$$

where μ is a trade-off parameter. Note that the tensor rank is not unique and there are several definitions to describe the tensor rank, including the Tucker rank, TT rank, and TR rank. Nonetheless, solving (11) using any of the above definitions suffers from intensive computational cost. As a result, it is essential to substitute a tractable surrogate for the rank term. In this manner, the tensor completion task can be further converted to solve

$$\min_{\mathcal{Z}} \Phi(\mathcal{Z}) + \frac{\mu}{2} \|\mathcal{P} \circ \mathcal{Z} - \mathcal{D}^{\text{obs}}\|_F^2 \quad (12)$$

where $\Phi(\mathcal{Z})$ represents a generalized low-rank constraint on \mathcal{Z} .

An effective way to solve this tough problem in (12) is performing the matrix factorization on the unfoldings of $\Phi(\mathcal{Z})$

$$\text{unfold}_n(\mathcal{Z}) = \mathbf{X}_n \mathbf{Y}_n \quad (13)$$

where $\text{unfold}_n(\mathcal{Z}) \in \mathbb{R}^{S_n \times T_n}$ has a predefined rank R_n , $\mathbf{X}_n \in \mathbb{R}^{S_n \times R_n}$, and $\mathbf{Y}_n \in \mathbb{R}^{R_n \times T_n}$. By inserting (13), the problem in (12) is transformed into the following Frobenius norm minimization problem:

$$\min_{\mathbf{X}, \mathbf{Y}, \mathcal{Z}} \frac{1}{2} \sum_{n=1}^L \|\mathbf{X}_n \mathbf{Y}_n - \text{unfold}_n(\mathcal{Z})\|_F^2 + \frac{\mu}{2} \|\mathcal{P} \circ \mathcal{Z} - \mathcal{D}^{\text{obs}}\|_F^2 \quad (14)$$

where L is the total number of unfolding matrices. $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_L)$ and $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_L)$. In the PMF algorithm [48], Tucker rank is adopted, and thus (14) can be rewritten as

$$\min_{\mathbf{X}, \mathbf{Y}, \mathcal{Z}} \frac{1}{2} \sum_{n=1}^N \|\mathbf{X}_n \mathbf{Y}_n - \mathbf{Z}_{(n)}\|_F^2 + \frac{\mu}{2} \|\mathcal{P} \circ \mathcal{Z} - \mathcal{D}^{\text{obs}}\|_F^2. \quad (15)$$

B. Seismic Reconstruction by TT Rank Optimization

The Tucker rank basically captures the correlation between one mode- k and the others. Fortunately, TT rank takes the correlation between the first n modes of \mathcal{D} and the rest of last $N - n$ modes of \mathcal{D} into account, thereby providing an more expressional manner to capture the global correlations of \mathcal{D} . As discussed in II-B, the TT rank can also be applied to define the tensor rank as a sum of the rank of mode- n canonical matricization of \mathcal{X} . Accordingly, the proposed PMF-TT simply replaces the Tucker rank with TT rank, and the formulation in (11) is changed to the following TT rank optimization problem:

$$\min_{\mathcal{Z}} \sum_{n=1}^L \text{rank}(\mathbf{Z}_{[n]}) + \frac{\mu}{2} \|\mathcal{P} \circ \mathcal{Z} - \mathcal{D}^{\text{obs}}\|_F^2 \quad (16)$$

where $L = N - 1$ because this unfolding approach generates $N - 1$ matrices. Similar to (14), the following equation is equivalent to (16):

$$\min_{\mathbf{X}, \mathbf{Y}, \mathcal{Z}} \frac{1}{2} \sum_{n=1}^{N-1} \|\mathbf{X}_n \mathbf{Y}_n - \mathbf{Z}_{[n]}\|_F^2 + \frac{\mu}{2} \|\mathcal{P} \circ \mathcal{Z} - \mathcal{D}^{\text{obs}}\|_F^2 \quad (17)$$

where $\mathbf{X}_n \in \mathbb{R}^{\prod_{j=1}^n I_j \times R_n}$ and $\mathbf{Y}_n \in \mathbb{R}^{R_n \times \prod_{j=n+1}^N I_j}$. Notably, (17) can be split into convex subproblems with respect to each block of the variables \mathbf{X}_n , \mathbf{Y}_n , and \mathcal{Z} , while the other two are fixed. Following PMF in [48], we use the block coordinate descent method to iteratively update different blocks of variables as follows.

1) *Update of \mathbf{X} and \mathbf{Y} :* The minimization subproblem with reference to \mathbf{X} and \mathbf{Y} can be calculated by

$$\min_{\mathbf{X}} \frac{1}{2} \sum_{n=1}^{N-1} \|\mathbf{X}_n \mathbf{Y}_n^k - \mathbf{Z}_{[n]}^k\|_F^2 \quad (18a)$$

$$\min_{\mathbf{Y}} \frac{1}{2} \sum_{n=1}^{N-1} \|\mathbf{X}_n^{k+1} \mathbf{Y}_n - \mathbf{Z}_{[n]}^k\|_F^2. \quad (18b)$$

For each \mathbf{X}_n and \mathbf{Y}_n ($n = 1, \dots, N - 1$), the optimization problem (18) can be solved by applying the following formula:

$$\mathbf{X}_n^{k+1} = \mathbf{Z}_{[n]}^k (\mathbf{Y}_n^k)^H \left(\mathbf{Y}_n^k (\mathbf{Y}_n^k)^H \right)^{\dagger} \quad (19a)$$

$$\mathbf{Y}_n^{k+1} = \left((\mathbf{X}_n^{k+1})^H \mathbf{X}_n^{k+1} \right)^{\dagger} (\mathbf{X}_n^{k+1})^H \mathbf{Z}_{[n]}^k \quad (19b)$$

where \dagger symbolizes the Moore–Penrose pseudoinverse. As shown in [48], (19b) can be simplified as follows:

$$\mathbf{X}_n^{k+1} = \mathbf{Z}_{[n]}^k (\mathbf{Y}_n^k)^H, \quad n = 1, \dots, N - 1. \quad (20)$$

It is rational to omit the computation of the Moore–Penrose pseudoinverse $\mathbf{Y}_n^k (\mathbf{Y}_n^k)^H$ because updating \mathcal{Z}^{k+1} in (22) merely needs the product of $\mathbf{X}_n^{k+1} \mathbf{Y}_n^{k+1}$ and the same results will be produced regardless of whether (19a) or (20) is used.

2) *Update of \mathcal{Z} :* Computing tensor \mathcal{Z} is related to addressing the following optimization subproblem:

$$\min_{\mathcal{Z}} \frac{1}{2} \sum_{n=1}^{N-1} \|\mathbf{X}_n^{k+1} \mathbf{Y}_n^{k+1} - \mathbf{Z}_{[n]}\|_F^2 + \frac{\mu}{2} \|\mathcal{P} \circ \mathcal{Z} - \mathcal{D}^{\text{obs}}\|_F^2. \quad (21)$$

By taking the partial derivative of \mathcal{X} in (21), the optimal closed-form solution of this subproblem can be obtained by

$$\mathcal{Z}^{k+1} = (\mathcal{I} - \alpha\mathcal{P}) \circ \mathcal{C} + \alpha\mathcal{D}^{\text{obs}} \quad (22)$$

with the parameter α specified by

$$\alpha = \frac{\mu}{N-1+\mu} \quad (23)$$

where the fourth-order tensor \mathcal{C} is determined by

$$\mathcal{C} = \frac{1}{N-1} \sum_{n=1}^{N-1} \text{fold}_{[n]}(\mathbf{X}_n^{k+1} \mathbf{Y}_n^{k+1}). \quad (24)$$

The pseudocode of the proposed PMF-TT is summarized in Algorithm 1.

Algorithm 1 Low-Rank Tensor Completion by PMF-TT

```

1: Inputs:  $\mathcal{D}^{\text{obs}}, \mathcal{P}, \mu, R_1, \dots, R_{N-1}$ 
2: Initializations:  $\mathbf{Y}^0, \mathcal{Z}^0, k = 0$ 
3: while the given stopping criterion is not satisfied do
4:   for  $n = 1, \dots, N-1$  do
5:     Unfolding  $\mathcal{Z}^k$  to get  $\mathbf{Z}_{[n]}$ 
6:      $\mathbf{X}_n^{k+1} = \mathbf{Z}_{[n]}^k (\mathbf{Y}_n^k)^H$ 
7:      $\mathbf{Y}_n^{k+1} = ((\mathbf{X}_n^{k+1})^H \mathbf{X}_n^{k+1})^\dagger (\mathbf{X}_n^{k+1})^H \mathbf{Z}_{[n]}^k$ 
8:   end for
9:    $\mathcal{C} = \frac{1}{N-1} \sum_{n=1}^{N-1} \text{fold}_{[n]}(\mathbf{X}_n^{k+1} \mathbf{Y}_n^{k+1})$ 
10:   $\mathcal{Z}^{k+1} = (\mathcal{I} - \alpha\mathcal{P}) \circ \mathcal{C} + \alpha\mathcal{D}^{\text{obs}}$ 
11:  set  $k = k + 1$ 
12: end while
13: Output:  $\mathcal{Z}^{k+1}$ 
```

C. Seismic Reconstruction by TR Rank Optimization

As discussed in Section II-C, the mode-1 and mode- $(N-1)$ canonical matricizations associated with TT rank are not well-balanced enough, therefore still leaving room for further development. Seeking to address this issue, we offer a novel model that takes advantage of the low TR rank structure by adopting a circular unfolding $\mathbf{Z}_{\{n,l\}}$, where $l = \lfloor (N/2) \rfloor$ for generating balanced matrices. As such, it captures the local correlation between floor $(N/2)$ modes and the rest ceil $(N/2)$ modes. By the new surrogate, we reformulate our model in (11) as follows:

$$\min_{\mathcal{Z}} \sum_{n=1}^L \text{rank}(\mathbf{Z}_{\{n,l\}}) + \frac{\mu}{2} \|\mathcal{P} \circ \mathcal{Z} - \mathcal{D}^{\text{obs}}\|_F^2 \quad (25)$$

where $L = N$ since this unfolding procedure produces N matrices. Similarly, an equivalent optimization problem can be derived

$$\min_{\mathbf{X}, \mathbf{Y}, \mathcal{Z}} \frac{1}{2} \sum_{n=1}^N \|\mathbf{X}_n \mathbf{Y}_n - \mathbf{Z}_{\{n,l\}}\|_F^2 + \frac{\mu}{2} \|\mathcal{P} \circ \mathcal{Z} - \mathcal{D}^{\text{obs}}\|_F^2 \quad (26)$$

and the following updates are easily accessed:

$$\mathbf{X}_n^{k+1} = \mathbf{Z}_{\{n,l\}}^k (\mathbf{Y}_n^k)^H \quad (27a)$$

$$\mathbf{Y}_n^{k+1} = ((\mathbf{X}_n^{k+1})^H \mathbf{X}_n^{k+1})^\dagger (\mathbf{X}_n^{k+1})^H \mathbf{Z}_{\{n,l\}}^k \quad (27b)$$

$$\mathcal{Z}^{k+1} = (\mathcal{I} - \alpha\mathcal{P}) \circ \mathcal{C} + \alpha\mathcal{D}^{\text{obs}}. \quad (27c)$$

We denote this new approach as PMF-TR, and Algorithm 2 presents the summarized pseudocode. The primary advantage of this algorithm is that circular unfolding produces more balanced matrices, which aids in the improvement of recovery performance.

Algorithm 2 Low-Rank Tensor Completion by PMF-TR

```

1: Inputs:  $\mathcal{D}^{\text{obs}}, \mathcal{P}, \mu, R_1, \dots, R_N$ 
2: Initializations:  $\mathbf{Y}^0, \mathcal{Z}^0, k = 0$ 
3: while the given stopping criterion is not satisfied do
4:   for  $n = 1, \dots, N$  do
5:     Unfolding  $\mathcal{Z}^k$  to get  $\mathbf{Z}_{\{n,l\}}$ 
6:      $\mathbf{X}_n^{k+1} = \mathbf{Z}_{\{n,l\}}^k (\mathbf{Y}_n^k)^H$ 
7:      $\mathbf{Y}_n^{k+1} = ((\mathbf{X}_n^{k+1})^H \mathbf{X}_n^{k+1})^\dagger (\mathbf{X}_n^{k+1})^H \mathbf{Z}_{\{n,l\}}^k$ 
8:   end for
9:    $\mathcal{C} = \frac{1}{N} \sum_{n=1}^N \text{fold}_{\{n,l\}}(\mathbf{X}_n^{k+1} \mathbf{Y}_n^{k+1})$ 
10:   $\mathcal{Z}^{k+1} = (\mathcal{I} - \alpha\mathcal{P}) \circ \mathcal{C} + \alpha\mathcal{D}^{\text{obs}}$ 
11:  set  $k = k + 1$ 
12: end while
13: Output:  $\mathcal{Z}^{k+1}$ 
```

On closer inspection of PMF-TR, and taking into consideration the fact that the observed seismic tensor \mathcal{D}^{obs} has exactly four dimensions ($N = 4$), we discover that $\mathbf{Z}_{\{n,l\}} = \mathbf{Z}_{\{n+l,l\}}^T$. In other words, the updating procedure in Algorithm 2 is redundant for $n = 3, 4$. To efficiently address the seismic reconstruction problem, we reduce half of the inner iterations regarding updating these three subproblems in Algorithm 2, and set $n = 1, 2$. In this particular case, the algorithm is designated as PMF-TRH and the corresponding pseudocode is presented in Algorithm 3. The only difference between PMF-TRH and PMF-TR is that the overall unfolding numbers of \mathcal{D}^{obs} in PMF-TRH ($n = 1, 2$) are less than that in PMF-TR ($n = 1, 2, 3, 4$).

Algorithm 3 Low-Rank Tensor Completion by PMF-TRH

```

1: Inputs:  $\mathcal{D}^{\text{obs}}, \mathcal{P}, \mu, R_1, R_2$ 
2: Initializations:  $\mathbf{Y}^0, \mathcal{Z}^0, k = 0$ 
3: while the given stopping criterion is not satisfied do
4:   for  $n = 1, 2$  do
5:     Unfolding  $\mathcal{Z}^k$  to get  $\mathbf{Z}_{\{n,l\}}$ 
6:      $\mathbf{X}_n^{k+1} = \mathbf{Z}_{\{n,l\}}^k (\mathbf{Y}_n^k)^H$ 
7:      $\mathbf{Y}_n^{k+1} = ((\mathbf{X}_n^{k+1})^H \mathbf{X}_n^{k+1})^\dagger (\mathbf{X}_n^{k+1})^H \mathbf{Z}_{\{n,l\}}^k$ 
8:   end for
9:    $\mathcal{C} = \frac{1}{2} \sum_{n=1}^2 \text{fold}_{\{n,l\}}(\mathbf{X}_n^{k+1} \mathbf{Y}_n^{k+1})$ 
10:   $\mathcal{Z}^{k+1} = (\mathcal{I} - \alpha\mathcal{P}) \circ \mathcal{C} + \alpha\mathcal{D}^{\text{obs}}$ 
11:  set  $k = k + 1$ 
12: end while
13: Output:  $\mathcal{Z}^{k+1}$ 
```

D. Accelerating the Proposed Methods With Sketching

The problem in (18) is indeed a low-rank matrix factorization problem. Consequently, many off-the-shelf algorithms, such as QR factorization and Cholesky decomposition, can also be implemented to update \mathbf{X} and \mathbf{Y} . We use the alternative

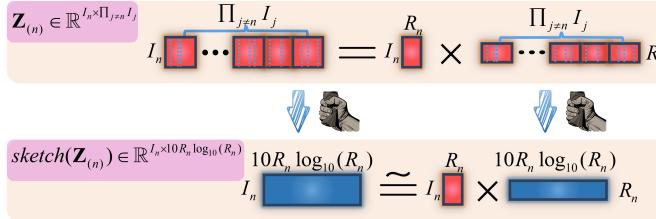


Fig. 3. Random sketching technique, using the mode- n unfolding of \mathbf{Z} as an illustration. This method enables an efficient solution to overdetermined least-squares problems by means of a randomized column sampling technique.

least squares in (19) and (27) since it is straightforward to implement and has high performance. In addition, recent studies [69] have demonstrated that randomized methods can accelerate this computation dramatically, which can be roughly divided into two categories: random projection-based methods and sampling-based methods. The former method projects any unfolding data $\text{unfold}_n(\mathbf{Z})$ or \mathbf{Z}_n into a considerably lower dimensional subspace by multiplying a random matrix. However, as a result of our test, this multiplication operation is also expensive since our proposed methods entail massive iterations. Moreover, constructing a random matrix for each iteration necessitates a significant increase in computational cost. The latter method performs a random sampling of certain rows or columns from the given matrix \mathbf{Z}_n , thus compressing \mathbf{Z}_n in this manner. Whether performing the sampling on the rows or columns depends on the size of $\mathbf{Z}_n \in \mathbb{R}^{S_n \times T_n}$. If $S_n \leq T_n$, solving \mathbf{X}_n is an over-determined process and we execute the sampling on the columns of \mathbf{Z}_n , and vice versa. Taking the original PMF as an example, one sampling method, namely, *sketching*, can be implemented by column slicing operations indexed by some random numbers as outlined in Algorithm 4, which has a considerable advantage over matrix multiplication in terms of running time. See Fig. 3 for a graphical illustration.

Algorithm 4 Sketching Sampling Algorithm

```

1: Inputs:  $\mathbf{Z}_n \in \mathbb{R}^{S_n \times T_n}$ ,  $\mathbf{Y}_n \in \mathbb{R}^{R_n \times T_n}$ ,  $R_n$ 
2: Outputs:  $\widehat{\mathbf{Z}}_n$ ,  $\widehat{\mathbf{Y}}_n$ 
3: Initializations: Sampling size  $SS_n = 10R_n \log_{10}(R_n)$   

    $\mathbb{P}(X = \alpha) = \frac{1}{T_n}$  for  $\alpha \in \{1, 2, \dots, T_n\}$ 
4: for  $r = 1, 2, \dots, SS_n$  do
5:   Pick  $t_r \in \{1, 2, \dots, T_n\}$  with  $\mathbb{P}$ 
6:   Set  $\widehat{\mathbf{Z}}_n(:, r) = \mathbf{Z}_n(:, t_r)$ 
7:   Set  $\widehat{\mathbf{Y}}_n(:, r) = \mathbf{Y}_n X(:, t_r)$ 
8: end for
```

We adopt the random sketching technique to reduce the running time of the PMF-based methods. In the case of the original PMF, the subproblem regarding solving \mathbf{X} can be reformulated as the following equation after applying a sketching technique:

$$\min_{\mathbf{X}} \frac{1}{2} \sum_{n=1}^N \|\mathbf{X}_n \text{sketch}(\mathbf{Y}_n^k) - \text{sketch}(\mathbf{Z}_{(n)}^k)\|_F^2 \quad (28)$$

where the updating of \mathbf{X} can be equivalently solved by a random least-squares problem. As the size of the new

least-squares subproblem is shrunk significantly, the computational complexity of updating the factor matrix is also reduced. We name this novel technique as PMF-sketch, and the entire procedure is illustrated in Algorithm 5. Similarly, we can also apply the sketching method to other PMF-based methods and obtain PMF-TT-sketch, PSMF-sketch, PMF-TR-sketch, and PMF-TRH-sketch. Note that sketching is used for updating \mathbf{Y} when $S_n > T_n$. In the following section of data examples, we will investigate their performance in detail.

Algorithm 5 Randomized Low-Rank Tensor Completion by PMF With Tucker Rank (PMF-Sketch)

```

1: Inputs:  $\mathcal{D}^{\text{obs}}$ ,  $\mathcal{P}$ ,  $\mu$ ,  $R_1, \dots, R_N$ 
2: Initializations:  $\mathbf{Y}^0, \mathbf{Z}^0$ ,  $k = 0$ 
3: while the given stopping criterion is not satisfied do
4:   for  $n = 1, \dots, N$  do
5:     Unfolding  $\mathbf{Z}^k$  to get  $\mathbf{Z}_{(n)}^k$ 
6:      $\mathbf{Y}_n^{k+1} = (\mathbf{X}_n^k)^H \mathbf{Z}_{(n)}^k$ 
7:      $\widehat{\mathbf{Z}}_{(n)}^k, \widehat{\mathbf{Y}}_n^{k+1} = \text{sketch}(\mathbf{Z}_{(n)}^k, \mathbf{Y}_n^{k+1}, R_n)$ 
8:      $\mathbf{X}_n^{k+1} = \widehat{\mathbf{Z}}_{(n)}^k (\widehat{\mathbf{Y}}_n^{k+1})^H (\widehat{\mathbf{Y}}_n^{k+1} (\widehat{\mathbf{Y}}_n^{k+1})^H)^\dagger$ 
9:   end for
10:   $\mathcal{C} = \frac{1}{N} \sum_{n=1}^N \text{fold}_n(\mathbf{X}_n^{k+1} \mathbf{Y}_n^{k+1})$ 
11:   $\mathbf{Z}^{k+1} = (\mathcal{I} - \alpha \mathcal{P}) \circ \mathcal{C} + \alpha \mathcal{D}^{\text{obs}}$ 
12:  set  $k = k + 1$ 
13: end while
14: Output:  $\mathbf{Z}^{k+1}$ 
```

We should supplement that the column or row selection procedure can be carried out using a variety of alternative probability distributions, as well as with or without replacement of data. Even though other distributions, such as leverage scores [70], might get better results, we use the uniform distribution here is because sampling other distributions needs more expensive computation.

E. Computational Complexity of Algorithms

Given a tensor \mathbf{Z} of order N , we assume $I_1 = \dots = I_{N-1} = I$ and set $R_1 = \dots = R_{N-1} = R$ for the analysis of complexity. The computational complexity of the PMF-based methods is shown in Table I. Note that the complexity for PSMF corresponds to $N = 4$ because it is not easy to square the unfolding for other values of N . We found that PMF-TRH presents the lowest complexity among the methods when sketching is not adopted. Sketching further reduces its computational complexity. Refer to Appendix A for the detailed derivation of computational complexity.

IV. EXPERIMENTS

A. Noise-Free Data Test

We first generate a 5-D noise-free seismic dataset to evaluate the computational efficiency and data reconstruction quality of the above-mentioned approaches. As depicted in Fig. 4(a), there are four curved events in this dataset, of which the size is $256 \times 20 \times 20 \times 10 \times 10$, comprising 20 traces in each common midpoint (CMP) dimension, ten traces in each offset dimension, and 256 time samples for each trace. A random

TABLE I
COMPUTATIONAL COMPLEXITY COMPARISON

Algorithm	Computational complexity
PMF	$\mathcal{O}(3NI^N R)$
PMF-sketch	$\mathcal{O}(2NI^N R)$
PMF-TT	$\mathcal{O}(3(N-1)I^N R)$
PMF-TT-sketch	$\mathcal{O}(2(N-1)I^N R)$
PSMF	$\mathcal{O}(3(N-1)I^N R)$
PSMF-sketch	$\mathcal{O}(2(N-1)I^N R)$
PMF-TRH	$\mathcal{O}(3(N-2)I^N R)$
PMF-TRH-sketch	$\mathcal{O}(2(N-2)I^N R)$
PMF-TR	$\mathcal{O}(3NI^N R)$
PMF-TR-sketch	$\mathcal{O}(2NI^N R)$

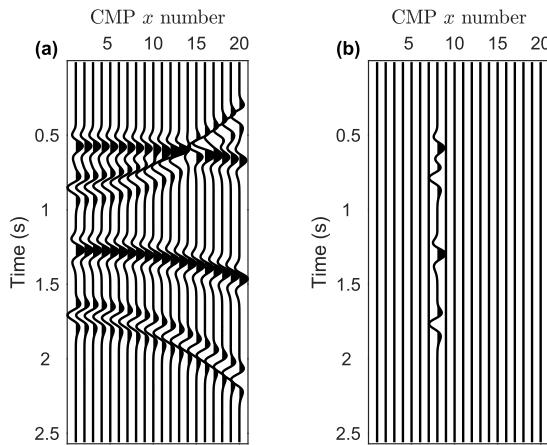


Fig. 4. (a) Slice view of noise-free synthetic complete data. (b) Decimated data with a missing rate of 90%.

amount of the total traces, such as 10%, 30%, 50%, 70%, and 90%, is removed from the ground-truth data to create the incomplete data. Thereafter, the proposed methods and the baseline methods, i.e., PMF and PSMF, are applied to complete each 4-D frequency cube. We set the reconstruction range of temporal frequencies to 1–70 Hz. Once all the frequency cubes have been processed, we assemble them into the 5-D tensor in order and then transform them into the original temporal–space domain. To allow for quantitative comparison, we introduce the reconstruction quality Q by

$$Q [\text{dB}] = 10 \log_{10} \left(\frac{\|\mathcal{D}^{\text{true}}\|_F^2}{\|\mathcal{D}^{\text{true}} - \mathcal{D}^{\text{recon}}\|_F^2} \right) \quad (29)$$

where $\mathcal{D}^{\text{true}}$ and $\mathcal{D}^{\text{recon}}$ represent the ground truth and reconstructed tensor in the time–spatial domain, respectively.

For noise-free testing, we set the same stop criteria for all methods. Specifically, there are two conditions for stopping the iteration: either it reaches a maximum iteration K_{iter} of 300 or the relative error is less than 10^{-4} , where the relative error is calculated by $((\|\mathcal{P} \circ \mathcal{Z}^{k+1} - \mathcal{D}^{\text{obs}}\|_F^2) / (\|\mathcal{D}^{\text{obs}}\|_F^2))$. The reinsertion parameter α regulates how much the observed data are reinserted into the final solution, and we set it to one since the observed data are not contaminated by noise. The last but most critical parameter pertains to the selection of an optimal rank for each method. For simplicity, we set $R = R_1 = \dots = R_{\max}$, where R_{\max} denotes the total unfolding

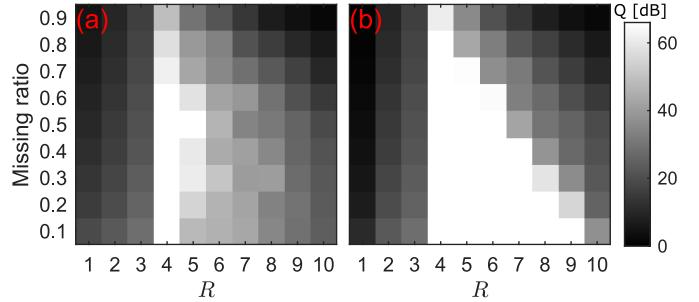


Fig. 5. Phase transition plot for (a) PMF and (b) PMF-sketch. After implementing the random sketching technique, PMF-sketch provides a substantially wider range of R values for high- Q reconstruction.

amount. For example, $R_{\max} = R_N$ in PMF, and $R_{\max} = R_{N-1}$ in PMF-TT. Then, we traverse all ranks from 1 to 12 with an increment of one to attain the optimal reconstruction quality for each method, as illustrated in Table II, where $\mathbf{R} = (R_1, \dots, R_{\max})$. As evident from green columns in the table, $R = 4$ tends to occur the most frequently, and therefore should be the optimal rank for most methods. Nevertheless, our proposed sketch-based methods sometimes can take the optimal rank $R = 7$, meaning that it has a higher R variance. In addition, Fig. 5 also demonstrates that implementation of the sketching technique broadens the optional range of R that could produce high- Q reconstruction, especially for small missing ratios. As a result, the sketching technique has the advantage of making R selection more feasible.

Then, we investigate the Q values in the red columns. On average, their reconstruction performance is similar since they can all acquire the best Q value for most cases. A closer inspection to 70% and 90% reveals that PMF-TRH performs best under no sketching conditions. Surprisingly, PMF-TT performs the worst, contrary to the findings in image recovery where PMF-TT should perform better than PMF and PSMF [71]. It is possible because the dimension of the fourth-order tensor is not large enough to fully exploit the advantages of TTs. In addition, we find that PSMF is still superior to PMF, although not by as much as in the original PSMF article [50]. This is most likely because the four spatial dimensions in this dataset are of different sizes. What is striking about Q values in this table is that almost every method is improved using the sketching approach and achieves a higher Q value. Specially, PSMF-sketch outperforms PSMF by 15.15 dB at 90% decimation. In addition, PMF-TT-sketch is also slightly superior to PSMF after the improvement gained by sketching. Consequently, the second advantage of the sketching technique is enhancing the reconstruction quality for noise-free data.

If we now turn to computational time in the orange columns, we can find that PMF has the longest run time, while PMF-TRH has the shortest run time in general. There are two possible explanations. In one regard, the unfolding scheme in PMF needs permutations, and it alters the data order in memory, therefore incurring additional costs. In contrast, PMF-TRH unfolds the tensor in a manner that only requires a moderate reshaping of the data without resorting the data in the computer memory. In another regard, PMF has

TABLE II
EVALUATION OF DIFFERENT PMF-BASED RECONSTRUCTION METHODS ON NOISY DATA

	Decimation 10%			Decimation 30%			Decimation 50%			Decimation 70%			Decimation 90%		
Method	Q	R	Time												
PMF	67.52	(4,4,4,4)	49.69	67.5	(4,4,4,4)	50.33	66.05	(5,5,5,5)	91.61	65.69	(4,4,4,4)	61.03	50.07	(4,4,4,4)	111.4
PMF-sketch	67.55	(6,6,6,6)	42.19	67.54	(4,4,4,4)	37.63	67.53	(4,4,4,4)	41.41	67.23	(4,4,4,4)	53.08	59.3	(4,4,4,4)	93.99
PMF-TT	67.48	(4,4,4)	39.48	67.51	(4,4,4)	34.39	67.49	(4,4,4)	39.8	50.99	(4,4,4)	62.72	40.62	(4,4,4)	70.61
PMF-TT-sketch	67.56	(6,6,6)	34.82	67.44	(5,5,5)	35.27	67.53	(4,4,4)	33.96	67.39	(4,4,4)	54.35	52.77	(4,4,4)	64.99
PSMF	67.43	(5,5,5)	21.03	67.52	(4,4,4)	27.09	67.52	(4,4,4)	23	67.25	(4,4,4)	26.5	52.15	(4,4,4)	45.33
PSMF-sketch	67.56	(5,5,5)	18.33	67.43	(7,7,7)	18.04	67.55	(5,5,5)	21.63	67.46	(4,4,4)	22.46	67.3	(4,4,4)	41.69
PMF-TRH	67.36	(4,4)	22.48	67.55	(5,5)	25.59	67.3	(5,5)	25.27	67.47	(4,4)	22.62	65.07	(4,4)	40.94
PMF-TRH-sketch	67.56	(4,4)	16.33	67.54	(4,4)	14.94	67.54	(4,4)	17.51	67.5	(4,4)	19.95	67.43	(4,4)	38.99
PMF-TR	67.25	(4,4,4,4)	24.72	67.46	(4,4,4,4)	27.02	66.92	(4,4,4,4)	27.56	58.52	(4,4,4,4)	33.09	61.77	(4,4,4,4)	63.91
PMF-TR-sketch	67.56	(7,7,7,7)	18.52	67.53	(6,6,6,6)	20.99	67.54	(4,4,4,4)	23.71	67.5	(4,4,4,4)	28.09	67.14	(4,4,4,4)	60.68

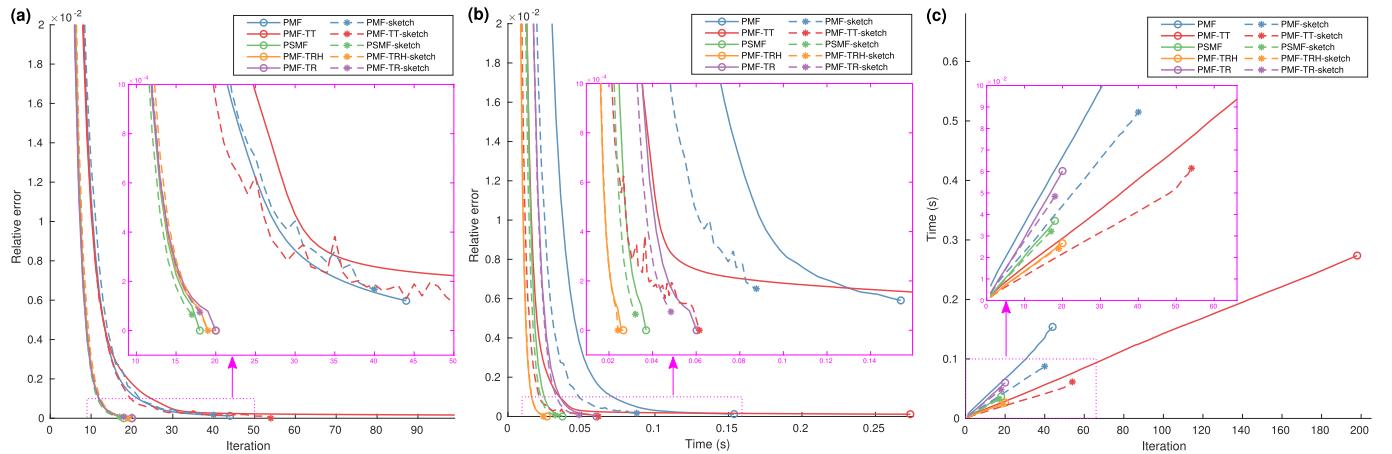


Fig. 6. Analysis of the recovery performance of PMF-based reconstruction methods for one frequency at 20 Hz under the 50% missing ratio. (a) Convergence versus iteration. (b) Convergence versus computational time. (c) Computational time versus iteration.

four unfoldings, whereas PMF-TRH unfolds the tensor only twice per iteration. In other words, PMF-TRH achieves better reconstruction performance with fewer unfoldings. In addition, we discover that PMF-TT consumes less time compared with PMF, despite having lower Q values. What can be clearly seen is the dramatic decline in run time after applying the sketching method. Although PMF-based methods are highly efficient, sketching can further reduce the computational cost by 40% in some instances and by more than 10% on average. Therefore, the third merit of the sketching strategy is to improve the efficiency of PMF-based methods.

To further explore the reason why sketching has higher Q values and shorter calculation time, we select a frequency cube at 20 Hz and repeat all methods 50 times on it with their optimal ranks to plot Fig. 6. Looking at Fig. 6(a), the convergence tendency of all methods is similar, where relative errors are decreasing dramatically, indicating that all methods are efficient. On closer examination of the amplified subfigure, we note that PMF and PMF-TT require more iterations to

reach the stop criteria. This is one reason they need more run time. Surprisingly, all horizontal positions of the asterisk markers are in front of their corresponding circle markers with the same color. These findings provide further evidence that sketching reduces the calculation time partly through a reduction in iteration numbers. Fig. 6(b) shows comparison of the execution time before and after sketching. End makers represent the average time taken to reach the stop criteria. From the above markers, we can conclude that applying the sketching takes less time to converge, which is consistent with our observations in Table II. Finally, we compare the calculation time of each iteration by measuring the slope of lines in Fig. 6(c). The smaller the slope, the shorter the time required for each iteration. It is evident that solid lines are always located above the dashed lines of the same color. Consequently, the sketching strategy accelerates the solution speed of each iteration for the overdetermined least-squares problem as we expect. Other than decreasing iterations, this provides an additional explanation as to why the implementation of

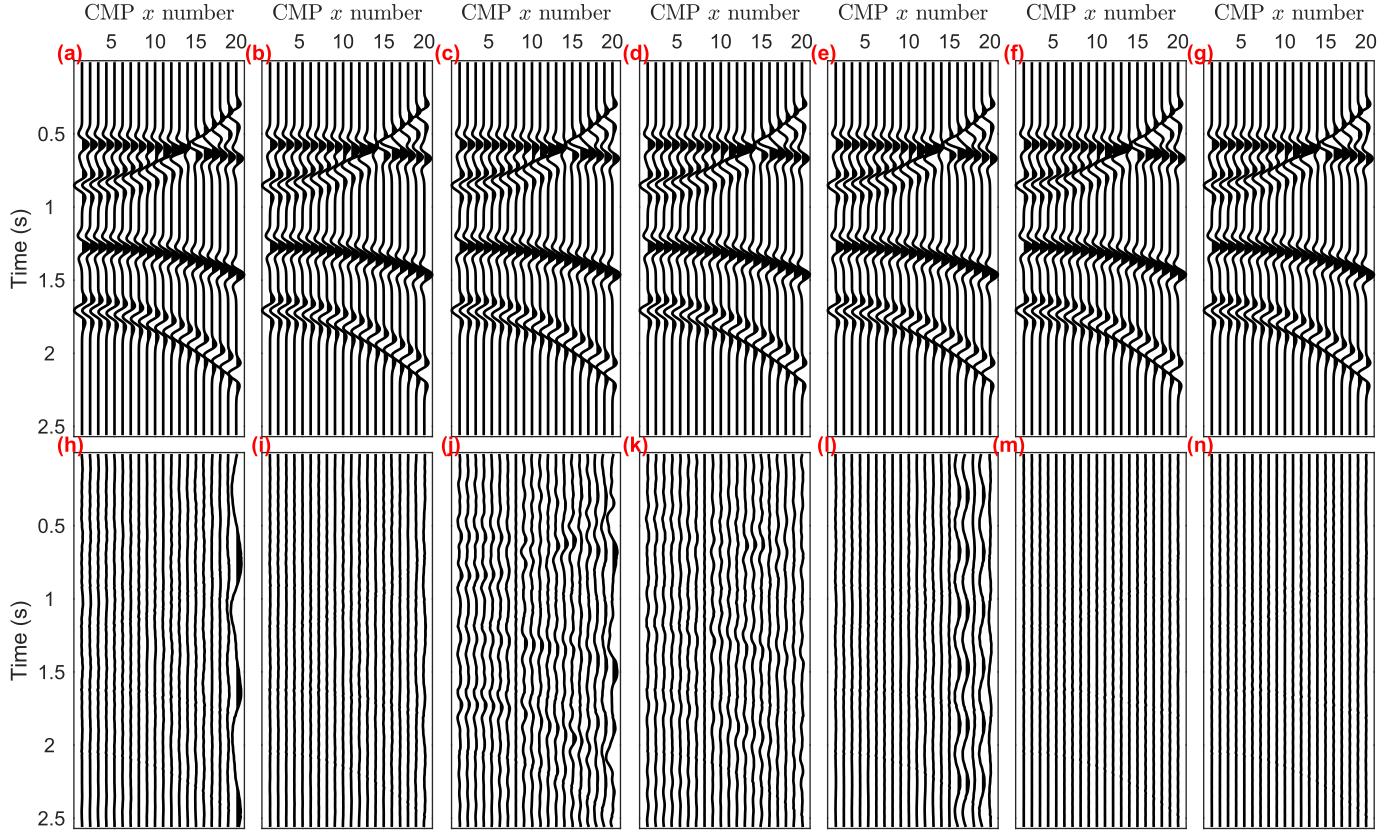


Fig. 7. Recover the noise-free data with 90% missing using (a) PMF, (b) PMF-sketch, (c) PMF-TT, (d) PMF-TT-sketch, (e) PSMF, (f) PSMF-sketch, and (g) PMF-TRH. (h)–(n) Corresponding reconstruction errors between (a)–(g) and complete data in Fig. 4(a), respectively. The data correspond to one slice of a 5-D volume. Error amplitudes are amplified by 100 times for better comparison.

sketching can reduce the total time costs. In addition, it can be observed that the slope of PMF-TT is not very large, indicating that overmuch iteration numbers are the main reason for the long run time of PMF-TT.

Finally, we visualize the reconstruction results of a 90% subsampling slice shown in Fig. 4(b) for more detailed examination. The last three methods shown in Table II are omitted from comparison because their performance is pretty much identical to PFM-TRH and their differences are hard to discern. Overall, it can be seen from Fig. 7(a)–(f) that all PMF-based methods achieve excellent performance and almost entirely recover the missing traces. To facilitate comparisons, we amplify the reconstruction errors between the reconstructed results and the true complete data by 100 times, as illustrated in Fig. 7(h)–(n). Without sketching, there are slight artifacts present in the error slice of PMF, PMF-TT, and PSMF. However, these artifacts appear much weaker after implementing the sketching method. These results corroborate our earlier findings of the superiority of the sketching technique. In addition, we can observe that the recovered traces by PMF-TRH are completely accurate, confirming that the proposed PMF-TRH is a promising alternative for seismic data reconstruction.

B. Noisy Data Test

We also evaluate the effectiveness of our proposed algorithms when earlier 5-D synthetic data are polluted by band-limited random noise, as plotted Fig. 8(a).

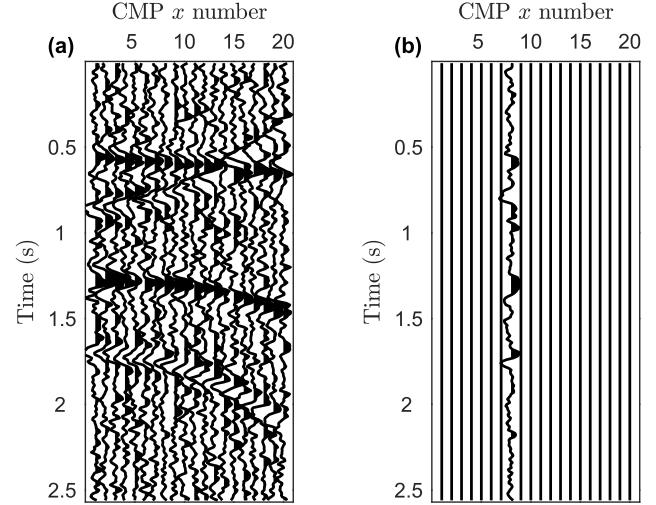


Fig. 8. (a) Slice view of noisy synthetic complete 5-D data. (b) Decimated data with a missing rate of 90%.

The signal-to-noise ratio of the corrupted data is 1.0 dB. Randomly removing 90% of the total traces from noisy data results in noisy and incomplete data, as shown in Fig. 8(b). In line with the previous noise-free experiment, we seek the best performance of each method by searching all available parameters. Note that the relative error for noisy data is computed in a different manner using $(\|\mathcal{Z}^{k+1} - \mathcal{Z}^k\|_F^2) / (\|\mathcal{Z}^k\|_F^2)$. The final parameter selection is summarized in Table III. It is

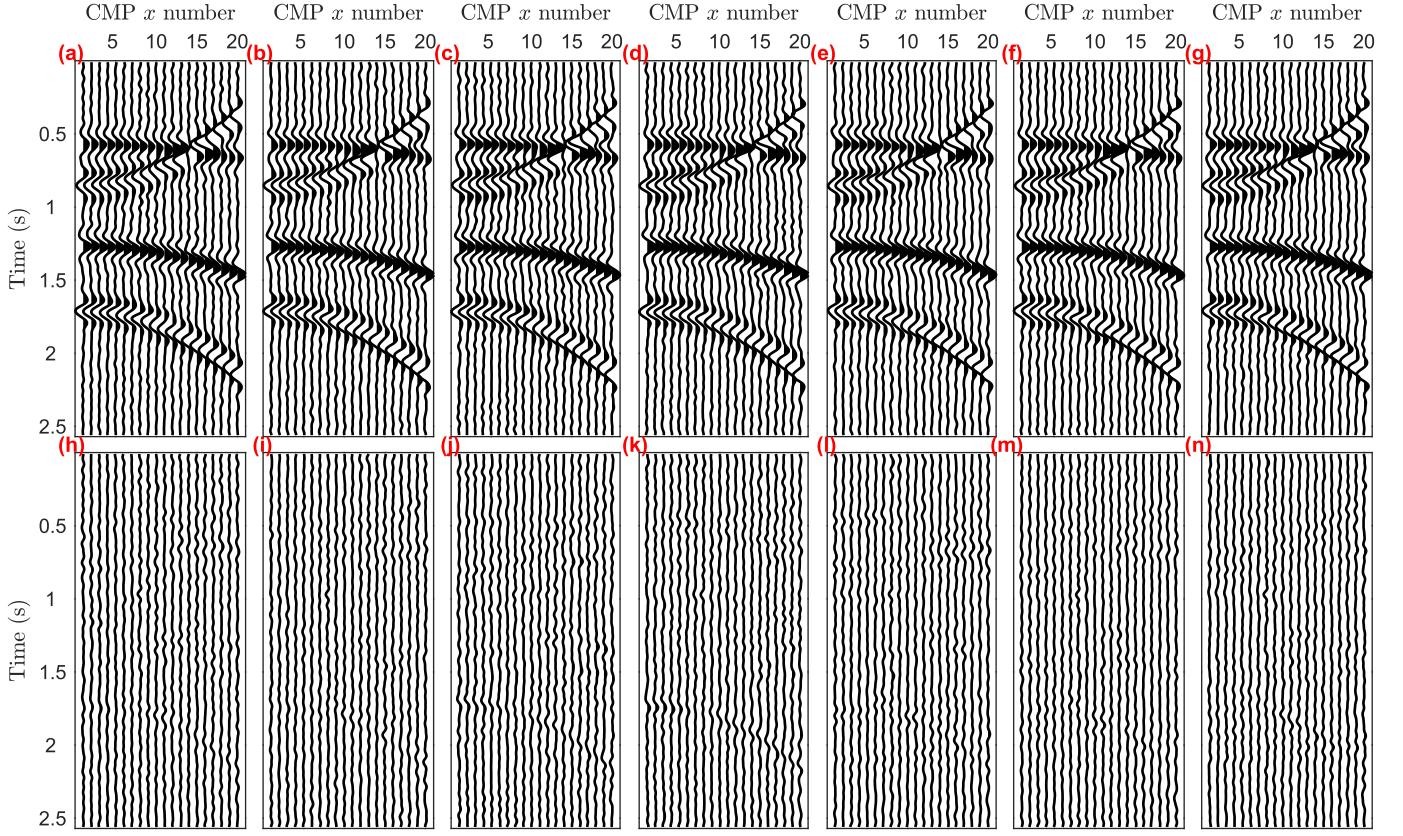


Fig. 9. Recover the noisy data with 90% missing using (a) PMF, (b) PMF-sketch, (c) PMF-TT, (d) PMF-TT-sketch, (e) PSMF, (f) PSMF-sketch, and (g) PMF-TR-sketch. (h)–(n) Corresponding reconstruction errors between (a)–(g) and complete data in Fig. 8(a), respectively. The data correspond to one slice of a 5-D volume.

apparent from these quantitative comparisons that adopting the sketching technique performs better in terms of both the evaluation measure Q and computational efficiency. Even without sketching, the proposed PMF-TRH methods can more accurately recover the tensor with less consuming time. PMF-TT performs worse than PMF but costs considerably less time because it also avoids time-consuming permutations. To further visually examine the performance, we display the reconstruction results in Fig. 9. From the figure, the superiority of the sketching method, both in the enhancement of reconstructed events and in the reduction of signal leakage, can be observed.

C. Real Field Data Example

Furthermore, we test the reliability of PMF-based reconstruction methods using a land dataset acquired in Canada. Fig. 10 presents the acquisition geometry. In the first step, we bin the data into regular grids that have a $20 \times 20 \text{ m}^2$ CMP area, a 400 m offset, and a 45° azimuth. Unlike synthetic tests, we use the offset $|h|$ and azimuth Az instead of h_x and h_y . Accordingly, we denote the incomplete 5-D real seismic dataset as $D^{\text{obs}}(\omega, x, y, |h|, Az)$, where $|h| = (h_x^2 + h_y^2)^{1/2}$ and $Az = \arctan((h_x/h_y))$. For any given bin, we average the traces falling within it. After the binning process, there is 96.79% preservation of the original traces in the grid and the binned data have 601 time samples, 38 CMP x points, 76 CMP y points, 12 offsets, and eight azimuths. The fold map, which

TABLE III
EVALUATION OF DIFFERENT PMF-BASED RECONSTRUCTION METHODS ON NOISY DATA

Method	Decimation 90%			
	Q	R	α	Time
PMF	10.39	(3,3,3,3)	0.2	124.91
PMF-sketch	10.53	(3,3,3,3)	0.2	104.7
PMF-TT	9.06	(3,3,3)	0.2	78.98
PMF-TT-sketch	9.41	(3,3,3)	0.2	66.53
PSMF	10.41	(4,4,4)	0.3	78.76
PSMF-sketch	11.04	(4,4,4)	0.3	73.82
PMF-TRH	10.43	(4,4)	0.3	62.37
PMF-TRH-sketch	10.5	(4,4)	0.3	58.71
PMF-TR	10.51	(4,4,4,4)	0.3	101.46
PMF-TR-sketch	10.67	(4,4,4,4)	0.3	97.99

indicates the number of traces per CMP, is depicted in Fig. 11 with a maximum fold of 15. Not surprisingly, this field example has a fairly large trace missing ratio of 94.39% since the average fold is only 5.38.

The entire survey data are divided into 405 blocks, and each block has a size of $100 \times 20 \times 20 \times 12 \times 8$. Overlapping ratios are

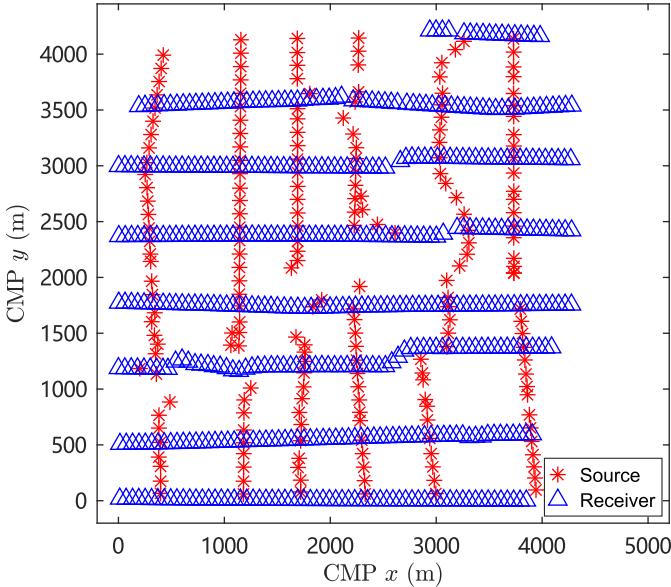


Fig. 10. Schematic of the seismic acquisition geometry of the field 5-D data.

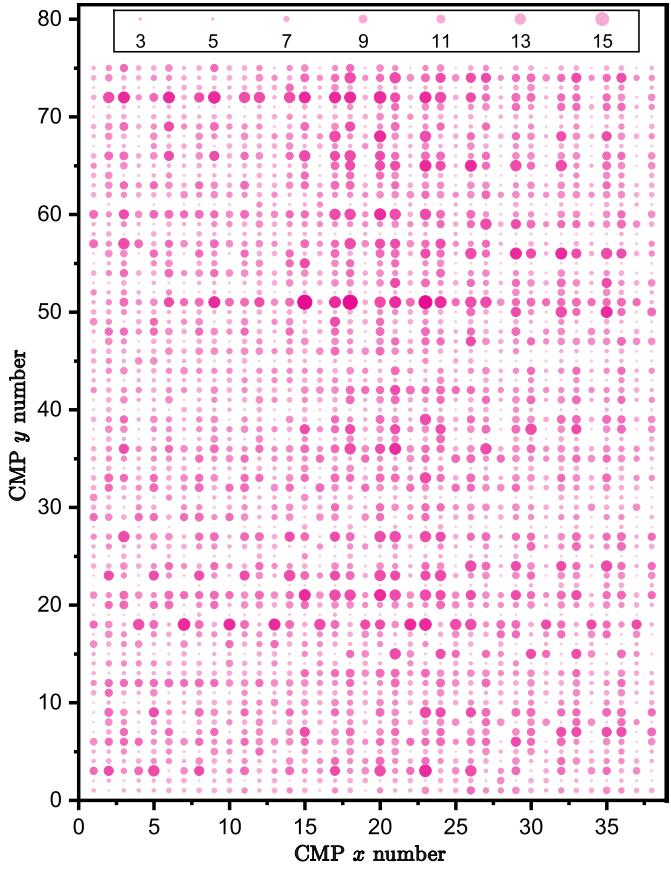


Fig. 11. Fold of coverage map for field 5-D data.

configured to 25% in time, CMP x , and CMP y dimensions. In the offset and azimuth directions, no overlap is applied since all bins in these two directions are completely occupied. Then, PMF-based algorithms are implemented at frequencies ranging from 1 to 65 Hz. By visually inspecting the quality of reconstruction, we ultimately set the same parameters for all methods: $R = 5$, $\alpha = 0.4$, and $K_{iter} = 300$. All programs

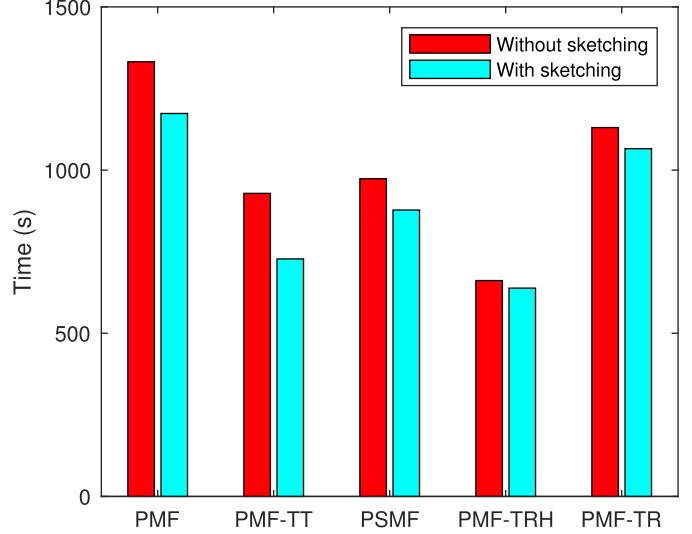


Fig. 12. Calculation time comparison of different reconstruction methods.

are run with MATLAB on the same server, which is equipped with 20 CPU cores of Intel(R) Xeon(R) CPU E5-2640 v4 at 2.40 GHz, 126 GB memory, and Ubuntu operating system. Fig. 12 displays the computation time required for different methods. It is evident that PMF-TT and PMF-TRH are the two most efficient methods before implementing the sketching approach. In addition, it is noteworthy in this figure that all computation time steadily declines after the implementation of sketching. These results are in accord with synthetic studies, in which using the randomized technique can shorten the computation time of the PMF-based reconstruction methods.

In the next step, we conduct the examination from the slice perspective. The synthetic example demonstrated that all the aforementioned methods are capable of delivering reconstruction results that are almost equally excellent. Accordingly, we solely present the results based on the PMF-TRH-sketch method since it is the most efficient method among the PMF-based methods.

Fig. 13 shows CMP y sections of the 5-D volume with the fixed CMP x (bin 4 and 60 m) and offset (bin 6 and 2000 m) prior to reconstruction. We can see that the seismic section of the original data in Fig. 13(a) is incomplete and contains considerable noise, presenting a significant challenge to successful reconstruction. By applying the PMF-TRH-sketch method, all the missing traces are completely recovered, as illustrated in Fig. 13(b), and these reflections are cleaner than those observed before reconstruction. Moreover, we can recognize that the continuity of seismic events is satisfactory since there is no substantial divergence between the original and reconstructed traces. Then, we examine CMP y slices of the 5-D volume with fixed offset (bin 6 and 2000 m) and azimuth (bin 4° and -45°) prior to reconstruction, as presented in Fig. 14(a). A large gap in incomplete observation complicates reconstruction. Fig. 14(b) provides the results obtained by the PMF-TRH-sketch method. It is clear that most of the missing data are accurately reconstructed, resulting in a much more coherent spatial representation of the seismic waveforms. Moreover, this technique succeeds in reconstructing the aforementioned big gap.

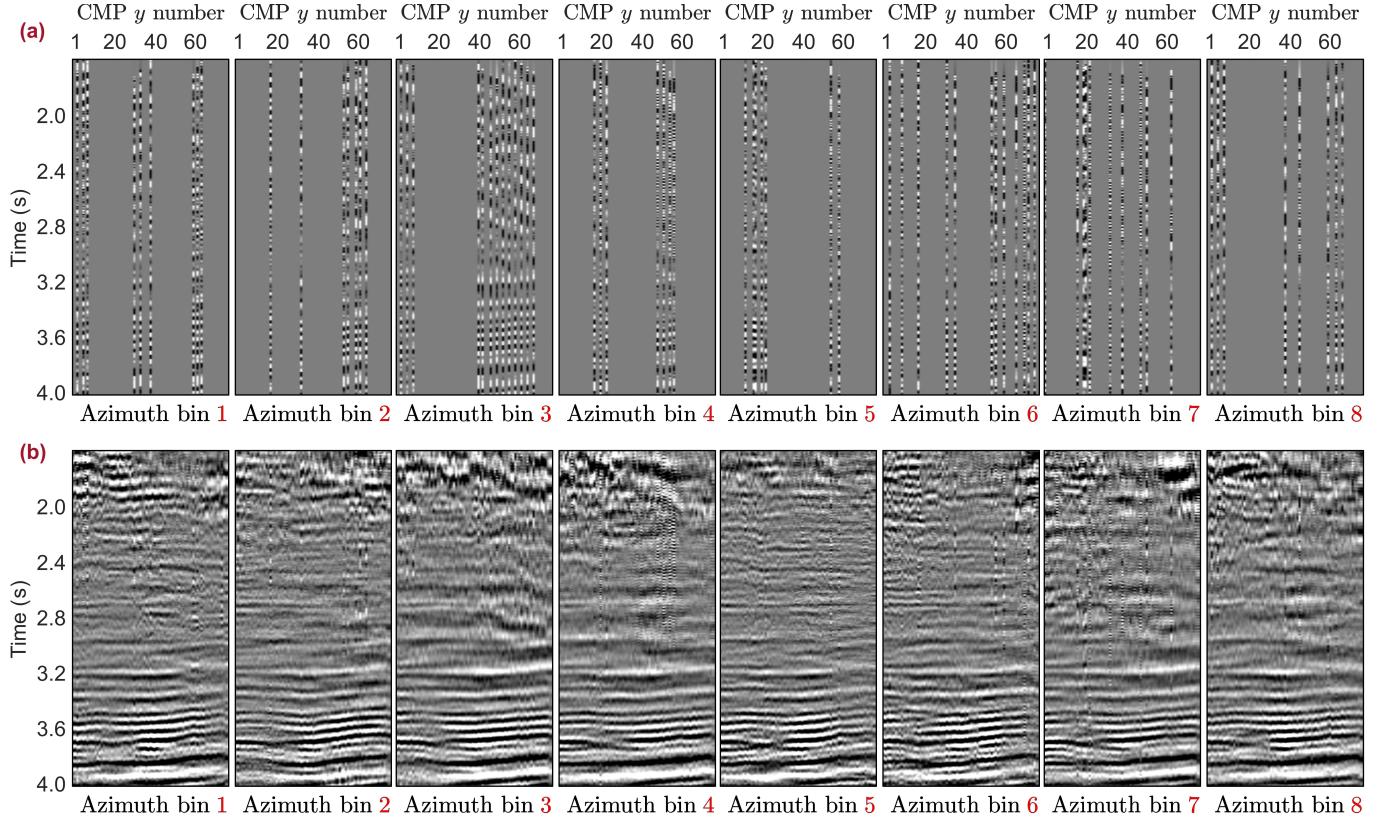


Fig. 13. Slice view of CMP y gather results (a) before and (b) after the reconstruction of the field data by assigning offset bin = 6 and CMP x number = 4. Azimuth bin [1, 8] is equivalent to $[-180^\circ, 135^\circ]$ with an increase of 45° .

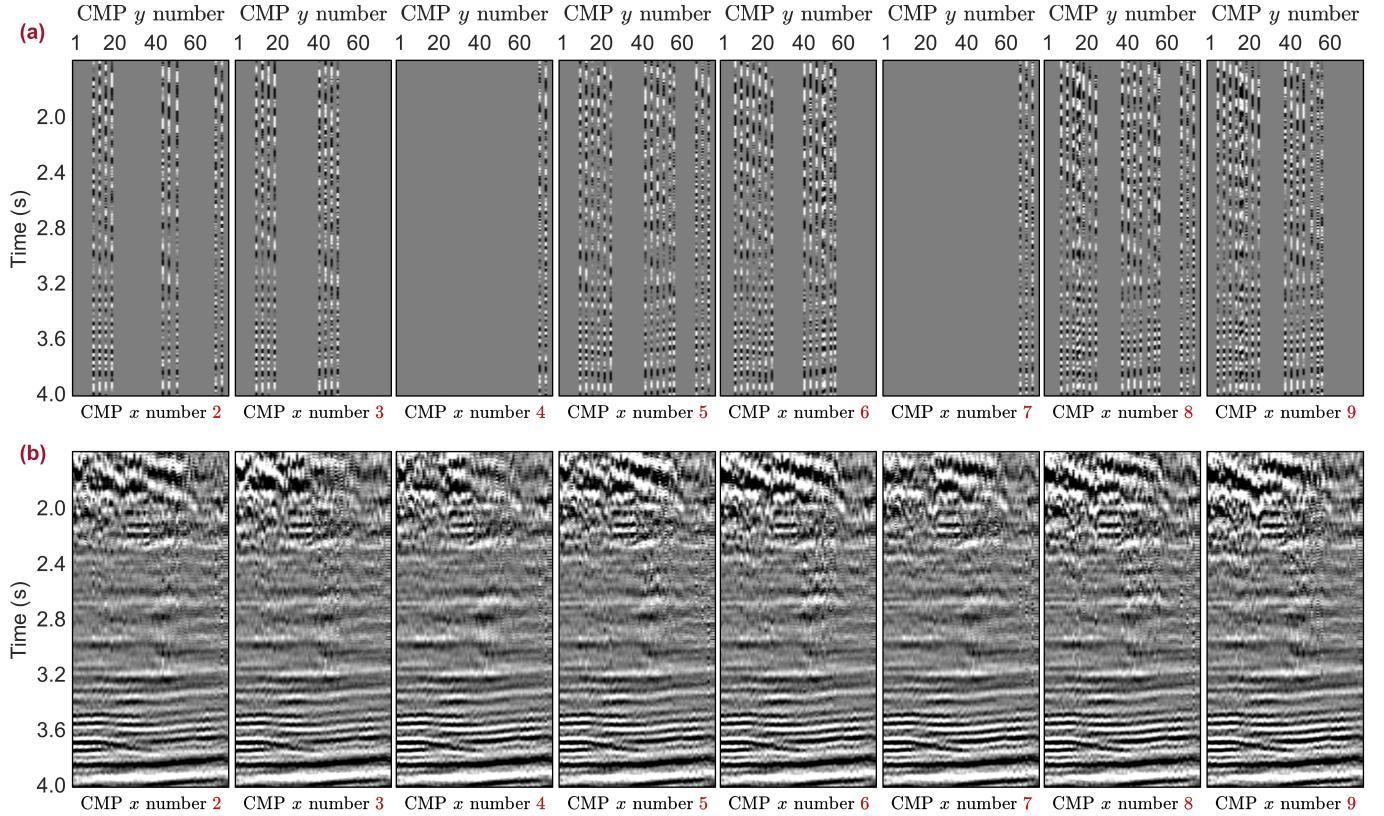


Fig. 14. Slice view of CMP y gather results (a) before and (b) after the reconstruction of the field data by assigning offset bin = 6 and azimuth bin = 4.

Fig. 15(a) illustrates the stacked data cube before interpolation. From this figure, we can see that the original stacked

data have a particularly low Q value, resulting in many reflections becoming obscured by noise. The stacked cube

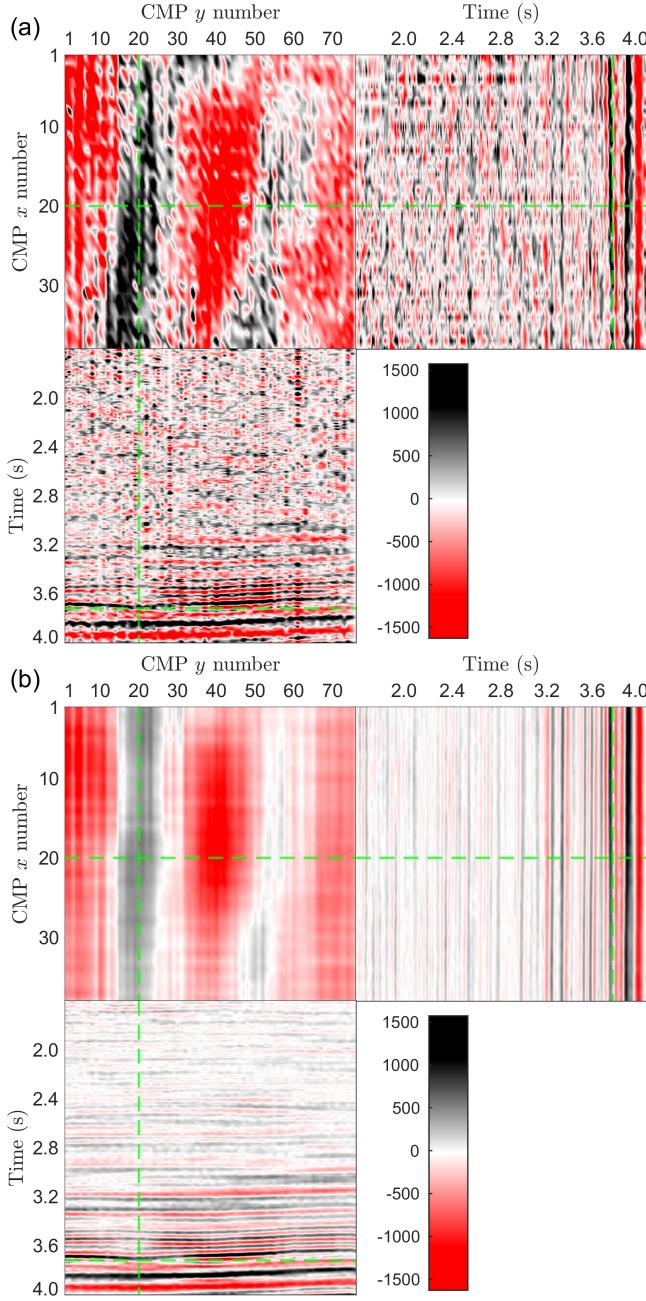


Fig. 15. Stacked results. (a) Before reconstruction. (b) After reconstruction using the PMF-TRH-sketch method.

after reconstruction is displayed in Fig. 15(b). From the comparison of the above two stacked results, the PMF-TRH-sketch eliminates prestack random noise effectively and significantly increases the signal energy. This result confirms the effectiveness of the PMF-TRH-sketch method in recovering noisy and irregularly sampled seismic data.

V. DISCUSSION

A. Rank Sensitivity Analysis

A priori knowledge of rank is required for rank-reduced tensor completion methods. The desired rank R for optimally reconstructing seismic data is often unknown. To simplify our analysis, in previous examples, we assumed all the rank values for TT and TR to be equal. In other words, $R_1 = \dots = R_{N-1}$

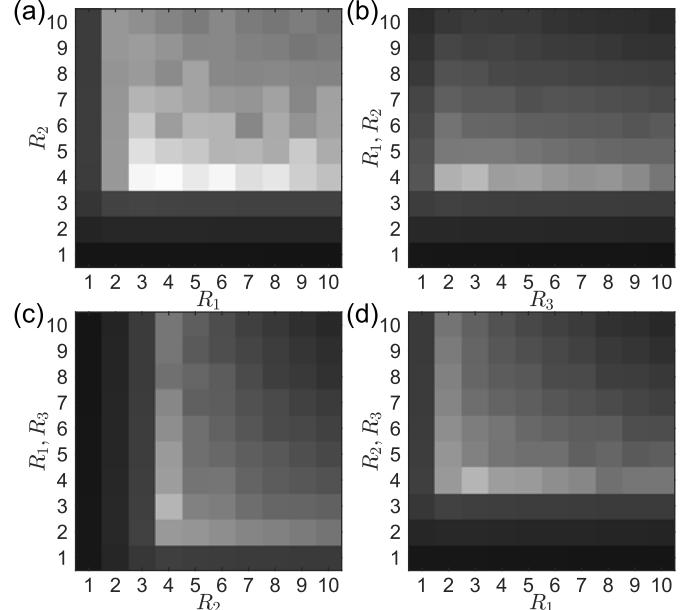


Fig. 16. Rank sensitivity analysis under the 90% decimation for (a) PMF-TRH and PMF-TT when (b) $R_1 = R_2$, (c) $R_1 = R_3$, and (d) $R_2 = R_3$. The colorbar is same as shown in Fig. 5.

for TT, $R_1 = \dots = R_N$ for TR, and $R_1 = R_2$ for TRH. In this section, we provide experiments that examined sensitivity to choosing different ranks for each unfolding. We restrict the analysis to PMF-TT and PMF-TRH.

This experiment follows the same settings as in Section IV-A, except for the assumption that the multiranks are identical. From Fig. 5, we observe that for moderate levels of decimations (e.g., 50%), a wide range of ranks provides accurate reconstructions; in contrast, for severe decimations of the order of 70%–90%, a more precise knowledge of the rank is needed. We also perform our experiment with a 90% decimation. From the PMF-TRH results in Fig. 16(a), we observe that satisfactory results are obtained with $R_1 \geq 1$ and $R_2 \geq 4$. Therefore, the reconstruction results are more sensitive to R_2 than R_1 . Rank R_2 corresponds to an unfolding matrix of size 200×200 , while R_1 associates with an unfolding of size 400×100 . Therefore, it is likely that the reconstruction results are more dependent on ranks corresponding to unfolding matrices, which more closely resemble square matrices. Similar conclusions can be drawn from the PMF-TT results in Fig. 16(b)–(d), where acceptable results require $R_2 \geq 4$ while merely demanding $R_1 \geq 1$ and $R_3 \geq 1$. This conclusion is consistent with previous articles stating that a well-balanced matricization scheme is vital for tensor completion [50], [71]. In addition, we observe that the globally optimal result is not far away from the optimal result obtained by the same multirank. For instance, the global optimal result for PMF-TT is $R_1 = 3$, $R_2 = 4$, and $R_3 = 3$, which is slightly better than the result obtained by $R_1 = R_2 = R_3 = 4$. Consequently, a practical method to find globally optimal results can entail two steps: first, search the suboptimal results with equal ranks; then, fine-tune the rank along each dimension with a small range of values.

B. Convergence Analysis

Our algorithms mainly comprise two parts. One is the PMF used to express the low rank of TT or TR decompositions, whose convergence has been demonstrated in [49]. The other one is the randomization method used to accelerate the computation of the low-rank matrix factorization, whose error bound has been derived in [72, Th. 4.10]. Following [69] and [73], we set the sampling size to $10R \log_{10} R$, whereas it lacks strict theoretical proof. Fig. 6 shows the convergence for the proposed methods. We have not demonstrated their convergence theoretically, but our experiments have served to heuristically show convergence for the type of practical problems studied in this article.

VI. CONCLUSION

Seismic reconstruction methods based on the TT rank and TR rank are proposed in this article. As demonstrated by synthetic experiments, both PMF-TT and PMF-TR can reconstruct 5-D seismic data with great accuracy. PMF-TT performs slightly worse than PMF in our datasets but requires less computation time. PMF-TR performs moderately better than PSMF but is more computationally intensive. After decreasing the half iteration chain of PMF-TR, PMF-TRH can achieve faster reconstruction without reducing the reconstruction quality. In addition, the randomized sketching technique is beneficial to all PMF-based approaches. Following that, we evaluate the performance of the most efficient method, PMF-TRH-sketch, on a real dataset. The field data results demonstrate that the PMF-TRH-sketch method not only successfully recovers missing traces in extremely incomplete data but also dramatically improves the quality of both prestack data and poststack data. Note that our intention is not to make comparisons among different PMF-based methods since they are all highly effective. By expanding the PMF-based reconstruction family, we aim to provide greater flexibility to practical implementation.

APPENDIX A

COMPUTATIONAL COMPLEXITY CALCULATION

We take PMF-TT as an example to calculate the computational complexity, but the same process can be used for the other methods explored in this article. As seen in (20) and (24), the updates of $\{\mathbf{X}_n^{k+1}\}_{n=1,\dots,N-1}$ and \mathbf{C} need a computational complexity of $\mathcal{O}((N-1)I^N R)$. In the update of \mathbf{Y}_n^{k+1} in (19b), the term $(\mathbf{X}_n^{k+1})^H \mathbf{X}_n^{k+1} \dagger$ requires computational complexity of $\mathcal{O}(I^n R^2 + R^3)$, and the other term $(\mathbf{X}_n^{k+1})^H \mathbf{Z}_{[n]}^k$ requires $\mathcal{O}(I^N R)$. Accordingly, the overall computational complexity of updating \mathbf{Y}_n^{k+1} can be summarized as $\mathcal{O}(I^N R + 2I^n R^2 + R^3)$. In fact, the low-rank nature of $\mathbf{Z}_{[n]} = \mathbf{X}_n \mathbf{Y}_n$, where $\mathbf{X}_n \in \mathbb{R}^{I^n \times R}$ and $\mathbf{Y}_n \in \mathbb{R}^{R \times I^{N-n}}$, often satisfies that $R \leq \min\{I^n, I^{N-n}\}$. Hence, $2I^n R^2 + R^3 \leq I^N R$, and the overall computational complexity of PMF-TT can be reduced to $\mathcal{O}(3(N-1)I^N R)$. After adopting the random sketching technique of a sampling size $10R \log_{10} R$, the sampling cost is $\mathcal{O}(10NR \log_{10} R)$ and can be ignored. The computational complexity of updating \mathbf{Y}_n^{k+1} can be reduced to $\mathcal{O}(10I^{N-n} R^2 \log_{10} R + 20R^3 \log_{10} R + R^3)$.

Similarly, since $R \leq \min\{I^n, I^{N-n}\}$, we can rewrite it as $\mathcal{O}(10I^{N-n} R^2 \log_{10} R)$. The overall computational complexity of PMF-TT-sketch is $\mathcal{O}(2(N-1)I^N R + 10[((2I(1 - I^{\lfloor ((N-1)/2 \rfloor})) / 1 - I) + I^{\lceil ((N-1)/2 \rceil)}]R^2 \log_{10} R)$ when N is even or $\mathcal{O}(2(N-1)I^N R + 10[((2I(1 - I^{\lfloor ((N-1)/2 \rfloor})) / 1 - I) + I^{\lceil ((N-1)/2 \rceil)}]R^2 \log_{10} R)$ when N is odd. When $N \geq 4$, we often have $I^{\lceil ((N-1)/2 \rceil)} \ll I^N$ and the leading order cost of PMF-TT-sketch can be roughly estimated by $\mathcal{O}(2(N-1)I^N R)$. In addition, a larger I and a smaller R lead to a closer approach to this value. Likewise, we can obtain the computational complexity of other methods and summarize them in Table I.

ACKNOWLEDGMENT

The authors Dawei Liu and Mauricio D. Sacchi would like to thank all the Signal Analysis and Imaging Group sponsors at the University of Alberta. They would also like to thank the comments from Jianjun Gao, which have improved the quality of this article.

REFERENCES

- [1] N. Kazemi, E. Bongajum, and M. D. Sacchi, "Surface-consistent sparse multichannel blind deconvolution of seismic signals," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3200–3207, Jun. 2016.
- [2] P. S. Schultz, "Seismic velocity estimation," *Proc. IEEE*, vol. 72, no. 10, pp. 1330–1339, Oct. 1984.
- [3] W. Zhang, J. Gao, Z. Gao, and H. Chen, "Adjoint-driven deep-learning seismic full-waveform inversion," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 10, pp. 8913–8932, Oct. 2021.
- [4] Y. Zhou, J. Gao, W. Chen, and P. Frossard, "Seismic simultaneous source separation via patchwise sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 9, pp. 5271–5284, Sep. 2016.
- [5] X. Wu, L. Liang, Y. Shi, and S. Fomel, "FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation," *Geophysics*, vol. 84, no. 3, pp. IM35–IM45, May 2019.
- [6] S. Spitz, "Seismic trace interpolation in the F-X domain," *Geophysics*, vol. 56, no. 6, pp. 785–794, Jun. 1991.
- [7] M. J. Porsani, "Seismic trace interpolation using half-step prediction filters," *Geophysics*, vol. 64, no. 5, pp. 1461–1467, 2012.
- [8] M. Naghizadeh and M. D. Sacchi, "F-X adaptive seismic-trace interpolation," *Geophysics*, vol. 74, no. 1, pp. V9–V16, Jan. 2009.
- [9] A. J. W. Duijndam, M. A. Schonewille, and C. O. H. Hindriks, "Reconstruction of band-limited signals, irregularly sampled along one spatial direction," *Geophysics*, vol. 64, no. 2, pp. 524–538, 1999.
- [10] M. A. Schonewille, R. Romijn, A. J. W. Duijndam, and L. Ongkiehong, "A general reconstruction scheme for dominant azimuth 3D seismic data," *Geophysics*, vol. 68, no. 6, pp. 2092–2105, Nov. 2003.
- [11] S. K. Chiu, "Multidimensional interpolation using a model-constrained minimum weighted norm interpolation," *Geophysics*, vol. 79, no. 5, pp. V191–V199, Sep. 2014.
- [12] M. D. Sacchi and T. J. Ulrych, "High-resolution velocity gathers and offset space reconstruction," *Geophysics*, vol. 60, no. 4, pp. 1169–1177, 1995.
- [13] M. D. Sacchi and T. J. Ulrych, "Estimation of the discrete Fourier transform, a linear inversion approach," *Geophysics*, vol. 61, no. 4, pp. 1128–1136, Jul. 1996.
- [14] M. D. Sacchi, T. J. Ulrych, and C. J. Walker, "Interpolation and extrapolation using a high-resolution discrete Fourier transform," *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 31–38, Jan. 1998.
- [15] P. Zwartjes and A. Gisolf, "Fourier reconstruction with sparse inversion," *Geophys. Prospecting*, vol. 55, no. 2, pp. 199–221, 2007, doi: [10.1111/j.1365-2478.2006.00580.x](https://doi.org/10.1111/j.1365-2478.2006.00580.x).
- [16] M. Naghizadeh and M. D. Sacchi, "Sparsity and band-limitation: Two sides of the same coin?" *Recorder*, vol. 38, no. 10, pp. 28–31, 2013.
- [17] B. Liu and M. D. Sacchi, "Minimum weighted norm interpolation of seismic records," *Geophysics*, vol. 69, no. 6, pp. 1560–1568, 2004.
- [18] D. Trad, "Five dimensional seismic data interpolation," in *Proc. SEG Tech. Program Expanded Abstr.*, 2008, pp. 978–982.

- [19] S. Jin, "5D seismic data regularization by a damped least-norm Fourier inversion," *Geophysics*, vol. 75, no. 6, pp. WB103–WB111, Nov. 2010.
- [20] V. Bardan, "Trace interpolation in seismic data processing," *Geophys. Prospecting*, vol. 35, no. 4, pp. 343–358, 1987.
- [21] M. N. Kabir and D. Verschuur, "Restoration of missing offsets by parabolic Radon transform," *Geophys. Prospecting*, vol. 43, no. 3, pp. 347–368, 1995.
- [22] W. Chen, L. Yang, H. Wang, and Y. Chen, "Fast high-resolution hyperbolic radon transform," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–10, 2022.
- [23] Z. Wang and Y. Li, "Trace interpolation using wavelet transform," in *Proc. SEG Tech. Program Expanded Abstr.*, 1994, pp. 729–730.
- [24] J. Liu, G. Liu, and Y. Chou, "Seismic data reconstruction via complex shearlet transform and block coordinate relaxation," *J. Seismic Explor.*, vol. 28, no. 4, pp. 307–332, 2019.
- [25] G. Hennenfent, L. Fenelon, and F. J. Herrmann, "Nonequispaced curvelet transform for seismic data reconstruction: A sparsity-promoting approach," *Geophysics*, vol. 75, no. 6, pp. WB203–WB210, Nov. 2010.
- [26] R. Shahidi, G. Tang, J. Ma, and F. J. Herrmann, "Application of randomized sampling schemes to curvelet-based sparsity-promoting seismic data recovery," *Geophys. Prospecting*, vol. 61, no. 5, pp. 973–997, Sep. 2013.
- [27] B. Wang, X. Chen, J. Li, and J. Cao, "An improved weighted projection onto convex sets method for seismic data interpolation and denoising," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 1, pp. 228–235, Jan. 2016.
- [28] L. Zhu, E. Liu, and J. H. McClellan, "Seismic data denoising through multiscale and sparsity-promoting dictionary learning," *Geophysics*, vol. 80, no. 6, pp. WD45–WD57, Nov. 2015.
- [29] H. Wang, W. Chen, Q. Zhang, X. Liu, S. Zu, and Y. Chen, "Fast dictionary learning for high-dimensional seismic reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 7098–7108, Aug. 2021.
- [30] S. Yu, J. Ma, X. Zhang, and M. D. Sacchi, "Interpolation and denoising of high-dimensional seismic data by learning a tight frame," *Geophysics*, vol. 80, no. 5, pp. V119–V132, Sep. 2015.
- [31] D. Liu, W. Wang, X. Wang, C. Wang, J. Pei, and W. Chen, "Poststack seismic data denoising based on 3-D convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 3, pp. 1598–1629, Nov. 2019.
- [32] H. Zhang, X. Yang, and J. Ma, "Can learning from natural image denoising be used for seismic data interpolation?" *Geophysics*, vol. 85, no. 4, pp. WA115–WA136, Jul. 2020.
- [33] D. A. B. Oliveira, R. S. Ferreira, R. Silva, and E. V. Brazil, "Interpolating seismic data with conditional generative adversarial networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 12, pp. 1952–1956, Dec. 2018.
- [34] B. Wang, N. Zhang, W. Lu, and J. Wang, "Deep-learning-based seismic data interpolation: A preliminary result," *Geophysics*, vol. 84, no. 1, pp. V11–V20, Jan. 2019.
- [35] F. Qian, Z. Liu, Y. Wang, S. Liao, S. Pan, and G. Hu, "DTAE: Deep tensor autoencoder for 3-D seismic data interpolation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–19, 2022.
- [36] M. J. Park, J. Jennings, B. Clapp, and B. Biondi, "Seismic data interpolation using a POCS-guided deep image prior," in *Proc. SEG Tech. Program Expanded Abstr.*, Sep. 2020, pp. 3154–3158.
- [37] P. Yuan *et al.*, "Self-supervised learning for anti-aliasing seismic data interpolation," in *Proc. 1st Int. Meeting for Appl. Geosci. Energy Expanded Abstr.*, Denver, CO, USA, Sep. 2021, pp. 1500–1504.
- [38] D. Liu, X. Yang, X. Wang, H. Mao, M. D. Sacchi, and W. Chen, "Deep learning for prestack strong scattered noise suppression," in *Proc. 1st Int. Meeting Appl. Geosci. Energy Expanded Abstr.*, Sep. 2021, pp. 1601–1605.
- [39] G. Ely, S. Aeron, N. Hao, and M. Kilmer, "5D and 4D pre-stack seismic data completion using tensor nuclear norm (TNN)," in *Proc. SEG Tech. Program Expanded Abstr.*, 2013, pp. 3639–3644.
- [40] C. D. Silva and F. Herrmann, "Hierarchical tucker tensor optimization—applications to 4d seismic data interpolation," in *Proc. 75th EAGE Expanded Abstr.*, 2013, p. 348.
- [41] S. Trickett, L. Burroughs, A. Milton, L. Walton, and R. Dack, "Rank-reduction-based trace interpolation," in *Proc. SEG Tech. Program Expanded Abstr.*, vol. 29, 2010, pp. 3829–3833.
- [42] V. Oropeza and M. Sacchi, "Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis," *Geophysics*, vol. 76, no. 3, pp. V25–V32, May 2011.
- [43] J. Gao, M. Sacchi, and X. Chen, "A fast reduced-rank interpolation method for prestack seismic volumes that depend on four spatial dimensions," *Geophysics*, vol. 78, no. 1, pp. V21–V30, 2013.
- [44] Y. Chen *et al.*, "Simultaneous denoising and reconstruction of 5-D seismic data via damped rank-reduction method," *Geophys. J. Int.*, vol. 206, no. 3, pp. 1695–1717, 2016.
- [45] N. Kreimer and M. D. Sacchi, "A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation," *Geophysics*, vol. 77, no. 3, pp. V113–V122, May 2012.
- [46] C. Da Silva and F. J. Herrmann, "Optimization on the hierarchical tucker manifold—Applications to tensor completion," *Linear Algebra Appl.*, vol. 481, pp. 131–173, Sep. 2015.
- [47] G. Ely, S. Aeron, N. Hao, and M. E. Kilmer, "5D seismic data completion and denoising using a novel class of tensor decompositions," *Geophysics*, vol. 80, no. 4, pp. V83–V95, Jul. 2015.
- [48] J. Gao, A. Stanton, and M. D. Sacchi, "Parallel matrix factorization algorithm and its application to 5D seismic reconstruction and denoising," *Geophysics*, vol. 80, no. 6, pp. V173–V187, Nov. 2015.
- [49] Y. Xu, R. Hao, W. Yin, and Z. Su, "Parallel matrix factorization for low-rank tensor completion," *Inverse Problems Imag.*, vol. 9, no. 2, p. 601, 2015.
- [50] J. Gao, J. Cheng, and M. D. Sacchi, "Five-dimensional seismic reconstruction using parallel square matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 4, pp. 2124–2135, Apr. 2017.
- [51] N. Lee and A. Cichocki, "Fundamental tensor operations for large-scale data analysis in tensor train formats," 2014, *arXiv:1405.7786*.
- [52] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, Jan. 2017.
- [53] Y. Zhang, Z. Han, and Y. Tang, "Color image denoising based on low-rank tensor train," in *Proc. 10th Int. Conf. Graph. Image Process. (ICGIP)*, H. Yu, Y. Pu, C. Li, and Z. Pan, Eds. Chengdu, China: SPIE, 2019, p. 41.
- [54] R. Dian and S. Li, "Hyperspectral image super-resolution via subspace-based low tensor multi-rank regularization," *IEEE Trans. Image Process.*, vol. 28, no. 10, pp. 5135–5146, Oct. 2019.
- [55] J. Xue, Y. Zhao, W. Liao, and J. C.-W. Chan, "Nonlocal low-rank regularized tensor decomposition for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 5174–5189, Jul. 2019.
- [56] J. Xue, Y. Zhao, W. Liao, J. C. Chan, and S. G. Kong, "Enhanced sparsity prior model for low-rank tensor completion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 11, pp. 4567–4581, Nov. 2020.
- [57] J. Xue, Y. Zhao, S. Huang, W. Liao, J. C.-W. Chan, and S. G. Kong, "Multilayer sparsity-based tensor decomposition for low-rank tensor completion," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jun. 18, 2021, doi: [10.1109/TNNLS.2021.3083931](https://doi.org/10.1109/TNNLS.2021.3083931).
- [58] I. V. Oseledets, "Tensor-train decomposition," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295–2317, Sep. 2011.
- [59] C.-Y. Ko, K. Batselier, L. Daniel, W. Yu, and N. Wong, "Fast and accurate tensor completion with total variation regularized tensor trains," *IEEE Trans. Image Process.*, vol. 29, pp. 6918–6931, 2020.
- [60] K. Xie *et al.*, "Neural tensor completion for accurate network monitoring," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Jul. 2020, pp. 1688–1697.
- [61] Z. Zhang, C. Ling, H. He, and L. Qi, "A tensor train approach for internet traffic data completion," *Ann. Oper. Res.*, 2021, doi: [10.1007/s10479-021-04147-4](https://doi.org/10.1007/s10479-021-04147-4).
- [62] Q. Zhao, G. Zhou, S. Xie, L. Zhang, and A. Cichocki, "Tensor ring decomposition," 2016, *arXiv:1606.05535*.
- [63] W. Wang, V. Aggarwal, and S. Aeron, "Efficient low rank tensor ring completion," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 5698–5706.
- [64] L. Yuan, Q. Zhao, and J. Cao, "Completion of high order tensor data with missing entries via tensor-train decomposition," 2017, *arXiv:1709.02641*.
- [65] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Sep. 2009.
- [66] Q. Zhao, M. Sugiyama, L. Yuan, and A. Cichocki, "Learning efficient tensor representations with ring-structured networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Brighton, U.K., May 2019, pp. 8608–8612.
- [67] S. Holtz, T. Rohwedder, and R. Schneider, "On manifolds of tensors of fixed TT-rank," *Numerische Math.*, vol. 120, no. 4, pp. 701–731, Apr. 2012.
- [68] J. Yu, G. Zhou, C. Li, Q. Zhao, and S. Xie, "Low tensor-ring rank completion by parallel matrix factorization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 7, pp. 3020–3033, Jul. 2021.

- [69] W. Gao and M. D. Sacchi, "Random noise attenuation via the randomized canonical polyadic decomposition," *Geophys. Prospecting*, vol. 68, no. 3, pp. 872–891, Mar. 2020.
- [70] P. Drineas, M. Magdon-Ismail, M. W. Mahoney, and D. P. Woodruff, "Fast approximation of matrix coherence and statistical leverage," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 3475–3506, Dec. 2012.
- [71] J. A. Bengua, H. N. Phien, H. D. Tuan, and M. N. Do, "Efficient tensor completion for color image and video recovery: Low-rank tensor train," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2466–2479, May 2017.
- [72] K. L. Clarkson and D. P. Woodruff, "Numerical linear algebra in the streaming model," in *Proc. 41st Annu. ACM Symp. Symp. Theory Comput. (STOC)*, Bethesda, MD, USA, 2009, p. 205.
- [73] C. Battaglino, G. Ballard, and T. G. Kolda, "A practical randomized CP tensor decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 39, no. 2, pp. 876–901, Jan. 2018.



Dawei Liu received the bachelor's degree in communication engineering from Chang'an University, Xi'an, China, in 2013, and the master's degree in electronic and communication engineering from Xi'an Jiaotong University (XJTU), Xi'an, in 2018, where he is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering.

His research interests include tensor decomposition, deep learning, and time–frequency analysis and their applications in seismic data processing.

Mr. Liu is a Reviewer of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.



Mauricio D. Sacchi (Member, IEEE) received the Diploma degree in geophysics from the National University of La Plata, La Plata, Argentina, in 1988, and the Ph.D. degree in geophysics from The University of British Columbia (UBC), Vancouver, BC, Canada, in 1996.

He joined the Department of Physics, University of Alberta, Edmonton, AB, Canada, in 1997. He has developed and taught short courses for industry and professional associations. He directs the Signal Analysis and Imaging Group, Edmonton, an initiative for advanced research for geophysical signal processing and imaging. His research interests include geophysical signal analysis and seismic imaging methods.

Dr. Sacchi was a recipient of the 2012 Medal of the Canadian Society of Exploration Geophysicists (CSEG). He was the 2014 Central and South America Honorary Lecturer for the Society of Exploration Geophysicists and the 2016 CSEG Distinguished Lecturer. He was the Editor-in-Chief of the *Geophysics* journal from 2016 to 2018. He was a recipient of the 2019 Virgil Kauffman Gold Medal.



Wenchao Chen received the B.S. and M.S. degrees in seismic exploration and information technology from Chang'an University, Xi'an, China, in 1993 and 1996, respectively, and the Ph.D. degree in electromagnetic field and microwave technology from Xi'an Jiaotong University, Xi'an, in 2000.

From 2000 to 2002, he was a Post-Doctoral Fellow with the Department of Computation Science, Northwestern Polytechnical University, Xi'an. Since 2002, he has been with the Institute of Wave and Information, Xi'an Jiaotong University, where he is currently a Professor. His research interests include seismic and ground-penetrating radar (GPR) signal processing, blind signal processing, and sparse representation.

Dr. Chen has served as a member for the Society of Exploration Geophysicists (SEG) and the Chinese Geophysical Society. He is a Referee for several journals, including the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, *Geophysics*, *Interpretations*, and *Journal of Applied Geophysics*.