Beating level-set methods for 3D seismic data interpolation: a primal-dual alternating approach

Rajiv Kumar[1], Oscar López[2], Damek Davis[3], Aleksandr Y. Aravkin[4] and Felix J. Herrmann[1]

[1]Department of Earth, Ocean and Atmospheric Sciences, The University of British Columbia

[2]Department of Mathematics, The University of British Columbia

[3] Department of Operations Research and Information Engineering, Cornell University

[4] Department of Applied Mathematics, University of Washington

Abstract

Acquisition cost is a crucial bottleneck for seismic workflows, and low-rank formulations for data interpolation allow practitioners to 'fill in' data volumes from critically subsampled data acquired in the field. Tremendous size of seismic data volumes required for seismic processing remains a major challenge for these techniques.

We propose a new approach to solve residual constrained formulations for interpolation. We represent the data volume using matrix factors, and build a block-coordinate algorithm with constrained convex subproblems that are solved with a primal-dual splitting scheme. The new approach is competitive with state of the art level-set algorithms that interchange the role of objectives with constraints. We use the new algorithm to successfully interpolate a large scale 5D seismic data volume, generated from the geologically complex synthetic 3D Compass velocity model, where 80% of the data has been removed.

Index Terms

Matrix completion, nuclear-norm relaxation, seismic data, seismic trace interpolation, alternating minimization, primal-dual splitting.

I. Introduction

Seismic data interpolation is crucial for accurate inversion and imaging procedures such as full-waveform inversion [32], reverse-time migration [4], [25] and multiple removal methods (SRME, EPSI) [31], [20]. Dense acquisition is prohibitively expensive in these applications, motivating reduction in seismic measurements. To show the scale of dense seismic data acquisition, we plot the synthetic 5D seismic data volume in Figure 1 at three levels of granularity. Note that, the details about this particular 5D seismic data are given in the experimental section. We are plotting seismic data volume in the canonical organization, where rows and columns of Figure 1 (a) represents (receiver-x, receiver-y) and (source-x, source-y) coordinates, respectively. This organization shows how we acquire seismic data in the field.

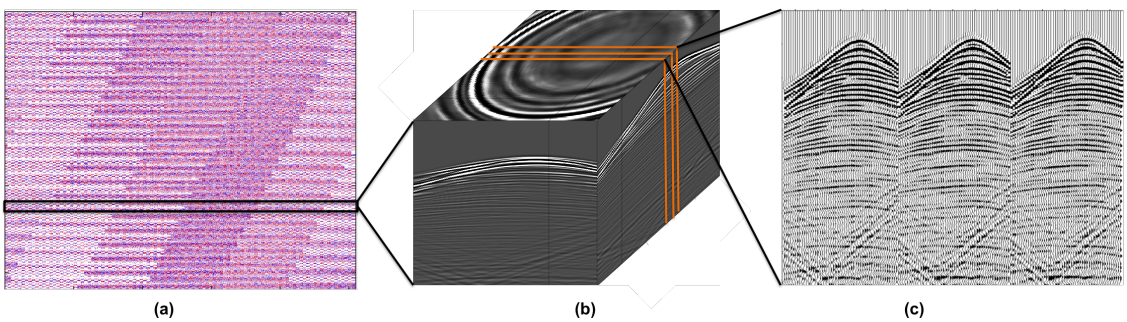


Fig. 1. To show the scale of the fully sampled 5D seismic data volume, we use three levels of granularity. (a) The first step shows a single 4D time-slice, consisting $101 \times 101 \times 40 \times 40$ iso-time samples, from the 5D tensor, where we use (s_x, s_y) matricization of the 4D tensor to display it as a matrix. The seismic data volume have 1024 time-slices. (b) In the second step, we extract **one** common-receiver gather at black box location from (a) where each common-receiver gather consists of $1024 \times 101 \times 101$ samples. (c) In the final step, for detailed visualization we extract 2D slices (marked with orange color) from (b) and unfold them along the source-axis. Interpolation is performed on the entire data volume, comprising multiple 4D monochromatic slices, each of containing $101 \times 101 \times 40 \times 40$ samples.

While subsampling reduces the acquisition cost, using subsampled sources and receivers without interpolation gives unwanted imaging artifacts. A range of trace interpolation methodologies have been proposed, exploiting low dimensional structure of seismic data, including sparsity ([14], [23]) and low-rank ([1], [18], [24], [29], [33]), and with theoretical guarantees for low-rank matrix recovery available in a range of contexts ([11], [26]). The main goal is to simultaneously sample and compress a signal using optimization to replace dense acquisition, thus enabling a range of applications in seismic data processing at a fraction of the cost.

Low-rank matrix completion techniques ([7], [6]) have been successfully applied to seismic trace interpolation. Here we focus on *residual-constrained* formulations, which minimize a regularizer subject to a constraint on the data misfit. This formulation is particularly appealing when practitioners have an estimate of the *noise floor* [30], [1].

In [1], the authors combine explicit low-rank factorization ([26], [28]) with *level-set optimization techniques* ([30], [3], [2]), and apply the resulting approach to seismic data interpolation in the midpoint-offset domain. A recent extension [22] includes seismic data sampled on unstructured grids, with recovery error bounds provided for the methodology.

Here, we develop a competitive optimization scheme based on alternating minimization for factorized formulations. The alternating approach in our context yields convex residual-constrained subproblems, which we solve using primal-dual splitting techniques of [8]. The resulting scheme is simple, and can be applied to extremely large-scale problems.

The paper proceeds as follows. In Section III, we formulate seismic data interpolation as a low-rank optimization problem, highlighting the transform domain that makes low-rank techniques applicable. In Section IV, we present the alternating optimization scheme for the residual-constrained formulation, together with an algorithm to solve the convex subproblems in each factor. The new approach is illustrated using 3D seismic data volumes in Section V, and we show it compares favorably to the the level-set approach of [1]. We end with a discussion and future directions in Section VI.

II. Notation

We use lower case to represent vectors (b, f), upper case to denote matrices and tensors (X, Y), and calligraphic upper case for linear operators ($\mathcal{A}$). For 3D seismic data acquisition, seismic data volumes have two source dimensions (s_x, s_y), two receiver dimensions (r_x, r_y), and time t.

III. Low-rank matrix completion

The goal of low-rank matrix completion is to accurately estimate the unobserved entries of a data matrix, $X \in \mathbb{C}^{n \times m}$, from the observed entries and prior knowledge that the matrix exhibits low-rank structure, i.e. has few nonzero singular values, or can be accurately approximated by

such a low-rank matrix. Letting $\Omega \subset \{1, 2, ..., n\} \times \{1, 2, ..., m\}$ be the set of observed entries, we define a sampling operator $P_\Omega : \mathbb{C}^{n\times m} \mapsto \mathbb{C}^{n\times m}$ by its element-wise action:

$$P_\Omega(X)_{ij} = \begin{cases} X_{ij}, & \text{if } \ (i,j) \in \Omega, \\ 0, & \text{otherwise.} \end{cases}$$

We can write our observations as $b = P_\Omega(X) + \epsilon$, where $\epsilon = P_\Omega(\epsilon) \in \mathbb{C}^{n\times m}$ models data corruption on the subset of observed entries, with a given noise floor $\|\epsilon\|_F \leq \eta$. Rather than minimizing the rank of X, which is a combinatorial problem, we can consider the nuclear norm $\|X\|_*$; in many contexts this allows exact recovery guarantees on X [11], [26]. This convex relaxation gives us the interpolated data volume $X^\sharp$ minimizer of

$$X^\sharp := \underset{X\in\mathbb{C}^{n\times m}}{\arg\min} \|X\|_* \ \text{ s.t. } \ \|P_\Omega(X) - b\|_F \leq \eta, \tag{1}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. This procedure finds the X with lowest nuclear norm that still fits observations up to the noise level η. In real-world applications, η may require estimation by cross-validation or other techniques.

Matrix completion via nuclear norm minimization has been extensively studied ([26], [7], [6], [11]). Typical results show that if Ω is generated uniformly randomly, then any $n\times m$ "incoherent" rank r matrix can be reconstructed with as few as $|\Omega| \sim \mathcal{O}(\max(n,m) r \operatorname{polylog} \max(n,m))$ observed entries. When $r \ll \min(n,m)$, this methodology provides a great reduction in the number of measurements needed compared to typical sampling procedures used in many applications. These insights have spurred work with the goal of solving (1) efficiently.

Many implementations to solve (1), such as singular-value projection and singular value thresholding ([15], [5]) require singular value decomposition (SVD) or partial SVD computations of the data matrix. Furthermore, treating the entire X as a decision variable requires storage and manipulation of an $n \times m$ object at every iteration. Storage and manipulation of large volumes becomes prohibitive for huge-scale problems, making (1) and equivalent formulations impractical for realistic seismic interpolations where the spatial interpolated grid for 3D seismic data acquisition is of the order 10^9. To rectify these issues ([1], [26], [28]) propose a matrix factorization approach. Stipulating a maximal rank r for X, the authors write the representation $X = LR^H$, with $L \in \mathbb{C}^{n\times r}$, $R \in \mathbb{C}^{m\times r}$ and ^{H} denoting the Hermitian transpose. As shown in ([26], [28]),

$$\|X\|_* = \min_{\{L,R:X=LR^H\}} \frac{1}{2}(\|L\|_F^2 + \|R\|_F^2) = \min_{\{L,R:X=LR^H\}} \frac{1}{2} \left\| \begin{bmatrix} L \\ R \end{bmatrix} \right\|_F^2,$$

motivating the formulation of [1]:

$$\min_{L\in\mathbb{C}^{n\times r},R\in\mathbb{C}^{m\times r}} \frac{1}{2}\left\|\begin{bmatrix}L\\R\end{bmatrix}\right\|_F^2 \quad \text{s.t.} \quad \|P_\Omega(LR^H)-b\|_F \leq \eta. \tag{2}$$

If $r \ll \min(n,m)$, we have reduced the memory requirements from mn to $rn+rm$, while projection onto the Frobenius norm ball is linear in the size of the variable, avoiding expensive SVD computations.

When X is a full-rank matrix that can be accurately approximated by a rank r matrix, implementing (2) with this choice of r provides reconstruction error bounds proportional to the error of the best r-rank approximation of X (see for example [6]). This holds true for many applications, where e.g. the data matrix is not low-rank but exhibits quickly decaying singular values, so that the error of the best r-rank approximation will be small for appropriately chosen r. When the underlying rank is not known, minimizing a regularization functional subject to a data constraint (2) has an important practical consequence: as the nominal rank r (number of columns in L and R) increases, we do not overfit the data [1].

Simply considering standard seismic data volumes does not immediately yield low-rank structure. In the next section, we show how seismic data can be transformed in order to find and exploit low-rank representation for interpolation.

A. *Low-rank structure of seismic data*

For 3D seismic data acquisition, each monochromatic tensor is a 4-dimensional volume with source dimensions s_x, s_y, and receiver dimensions r_x, r_y, respectively. In order to exploit low-rank structure for 3D seismic data acquisition, [10], [19] proposed two choices of *matricization*, or operation of unfolding a tensor into a matrix along specific dimensions. We either place the r_x, r_y dimensions in the rows and s_x, s_y dimensions in the columns, or r_y, s_y dimensions in the rows and r_x, s_x dimensions in the columns. In r_x, r_y matricization (Figure 1 (a)), operator P_Ω samples observed rows and/or columns, i.e. observed data corresponding to source and/or receiver combinations as shown in Figure 2 (a). Placing both receiver coordinates along the rows, i.e., matricization r_x, r_y, yields a matrix that has high rank (Figure 3a, red curve) and action of the subsampling operator (in particular removal of rows and/ or columns) can only *decrease* the rank (Figure 3b, red curve). Therefore, interpolation of the data by rank penalization is doomed to

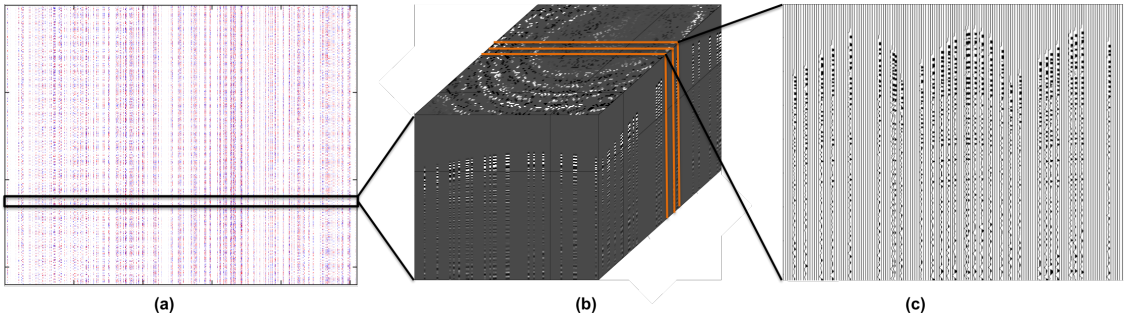
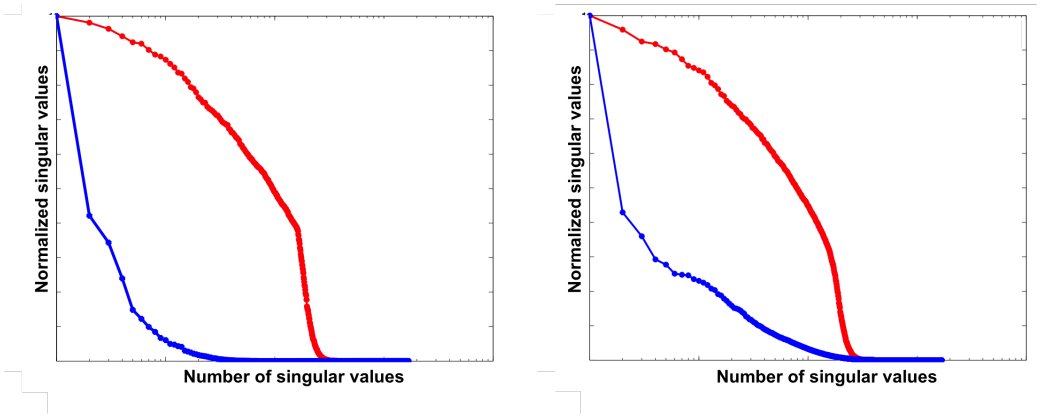


Fig. 2. Same as Figure 1 but after removing 80% jittered sources. Our goal is to recover fully sampled data as shown in Figure 1 from 80% subsampled data.

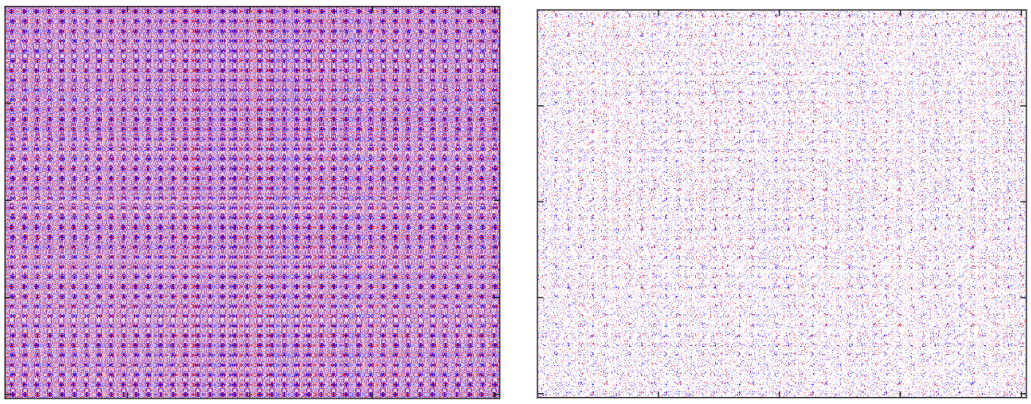
failure when using r_x, r_y matricization, since the fully sampled object has higher rank (and larger singular values) than the subsampled dataset. We therefore look for a transform domain where the fully sampled data matrix exhibits low-rank structure, while subsampling by P_Ω increases the rank and/or negatively alters the fast decay of singular values. In this context, we can expect low-rank penalization to help. We use the rank-revealing transforms of [1], [9], [19]. To motivate the use of these transforms, we first explore the singular value decay of full and subsampled volumes in the context of 3D seismic data acquisition. Note that, monochromatic seismic data matrices are full rank even in the transformed domain, but they have fast decay of the singular values in the transform domain, which means that these matrices can be well approximated with low-rank matrices. Therefore, we assume that the seismic data matrix exhibits low-rank structure when it has fast-decay of singular values and high-rank structure when it has slow-decay of singular values. In case of 3D seismic data acquisition, matricization r_x, s_x (Figure 4 (a)) yields fast decay of the singular values (Figure 3a, blue curve) for the original fully sampled data volume, while subsampling (Figure 4 (b)) causes the singular values to decay at a slower rate (Figure 3b, blue curve). Therefore, we select the latter matricization for 3D seismic data acquisition. Note that, we are only showing the 4D slice of a fully sampled and subsampled data volume for r_x, s_x matricization in Figure 4. The representation of a common-receiver for full and subsampled data is already shown in Figure 1 (b,c).

Working under an appropriate rank-revealing transformation domain, the subsampling operator P_Ω will also involve a transformation operator represented by $\mathcal{S} : \mathbb{C}^{n\times m} \mapsto \mathbb{C}^{p\times q}$. This transformation operator projects the seismic data from the source-receiver domain to the rank-revealing transform domains and its adjoint reverse the operation. We call this sampling-transformation



(a) (b)

Fig. 3. Singular value decay of (s_x, s_y) matricization (red curve) and (r_x, s_x) matricization (blue curve). (a) Fully sampled data and (b) subsampled data.



(a) (b)

Fig. 4. Same as Figure 1 but in (r_x, s_x) matricization. (a) Fully sampled and (b) subsampled data after removing 80% jittered sources. We are only showing the 4D-slice here because the representation of common-receiver gather will remain same as shown in Figures 1,2.

operation $\mathcal{A}(\cdot) = P_\Omega \mathcal{S}^T(\cdot) : \mathbb{C}^{p\times q} \mapsto \mathbb{C}^{n\times m}$, where ^{T} represents the adjoint, and minimize (2) with P_Ω replaced by $\mathcal{A}$. Note that, for interpolation, we permute the dimensions of seismic data while going from r_x, r_y matricization to r_x, s_x matricization, hence, overall number of samples in a monochromatic 4D tensor remains the same.

IV. Methodology

While several formulations are available to solve the equality constrained version of (2) (i.e. where $\eta = 0$), see [26], and penalized formulations, e.g. [28], few focus on the case where η is

provided by the user. The level-set factorized approach, which we call LR-BPDN following [1], solves (2) for a prescribed η by defining the value function

$$v(\tau) = \min_{\|L\|_F^2 + \|R\|_F^2 \leq 2\tau} \|\mathcal{A}(LR^H) - b\|_F,$$

and applying Newton's method to find $v(\tau) = \eta$. In practice this means inexactly solving a sequence of optimization problems to evaluate $v(\tau)$, and using duality theory to compute $v'(\tau)$ in order to compute the next τ iterate [30], [1]. The overall approach is reliable but can take many iterations, especially for high-fidelity data fitting.

Here, we propose a block-coordinate descent scheme, outlined in Algorithm 1.

Algorithm 1 Nuclear Norm Minimization via Alternating Optimization

Require: $\mathcal{A}$, b, r, η, K_0

1: **initialize:** any $L^0 \in \mathbb{C}^{p\times r}$, $\eta_0 = \|b\|_F$, $\alpha = 0.1$.

2: **for** $k = 0$ to $K_0 - 1$ **do**

3: $\eta_{k+1} = \max(\alpha^k \eta_k, \eta)$.

4: $R^{k+1} = \underset{\{R\in\mathbb{C}^{q\times r}\}}{\arg\min} \frac{1}{2}\|R\|_F^2$ s.t. $\|\mathcal{A}(L^k R^H) - b\|_F \leq \eta_{k+1}$

5: $L^{k+1} = \underset{\{L\in\mathbb{C}^{p\times r}\}}{\arg\min} \frac{1}{2}\|L\|_F^2$ s.t. $\|\mathcal{A}(LR^{H^{k+1}}) - b\|_F \leq \eta_{k+1}$

6: **end for**

7: **return** $(X^\sharp := \mathcal{S}^T(L^{K_0}\left(R^H\right)^{K_0}))$

Steps 3 and 4 focus on a single matrix factor at a time. Alternating approaches are common for matrix factorization ([21], [16], [12]). The main competing algorithms for constrained matrix completion formulations (2) use a level-set approach along the lines of [1] (see e.g [21]).

We implement steps 3 and 4 using a matrix-free approach, described in detail below.

A. Solving Steps 3 and 4 with Primal-Dual Splitting

All of the work of Algorithm 1 occurs in steps 3 and 4. Each step is an inequality constrained, convex optimization problem. In typical primal-only iterative optimization algorithms, such as the projected-gradient method, we must, at every iteration, project onto the constraint set. In large-scale problems, such a projection itself requires an iterative algorithm. Instead, we solve steps 3 and 4 using the *primal-dual* approach developed in [8], each step of which consists solely of matrix-vector products and scalar multiplication, and which is completely free of projections,

making it suitable for extremely large-scale problems. Moreover, it is very simple to implement: the full details of the approach are contained in Algorithm 2 for the L-update (step 5). The R-update (step 4) is obtained immediately by changing the roles of L and R. The matrix norm $\|\cdot\|_{\mathrm{op}}$ used by Algorithm 2 for stepsize selection is the largest singular value of the input matrix.

Algorithm 2 Primal-Dual Splitting Algorithm for Step 5 in Algorithm 1

Require: $\mathcal{A}$, b, r, η, K_1, $R \in \mathbb{C}^{q\times r}$

1: **initialize:** $k = 0$, any $y^0 \in \mathbb{C}^{n\times m}, L^0 \in \mathbb{C}^{p\times r}$

2: $\quad \mathcal{A} : \mathbb{C}^{p\times r} \to \mathbb{C}^{n\times m}$ defined by $L \mapsto \mathcal{A}(LR^H)$

3: $\quad\quad \gamma = \frac{0.99}{\|R\|_{\mathrm{op}}}$

4: **for** $k = 0$ to $K_1 - 1$ **do**

5: $\quad L^{k+1} = \frac{1}{1+\gamma}(L^k - \gamma\mathcal{A}^T y^k)$

6: $\quad y^+ = y^k + \gamma\mathcal{A}(2L^{k+1} - L^k) - \gamma b$

7: $\quad y^{k+1} = \max\left\{1 - \frac{\eta\gamma}{\|y^+\|_F}, 0\right\} y^+.$

8: **end for**

9: **return** $(L = L^{K_1})$

To understand Algorithm 2, we formulate a convex-concave saddle-point representation of step 5 of Algorithm 1:

$$\min_L \max_y \left\{\frac{1}{2}\|L\|^2 + \langle \mathcal{A}L - b, y\rangle - \eta\|y\|_2\right\}. \tag{3}$$

Computing the maximum over y in (3) immediately recovers the primal problem in L for step 5 of Algorithm 1, since this maximum is equal to $\frac{1}{2}\|L\|^2$ when $\|\mathcal{A}L - b\| \leq \eta$, and is infinity otherwise. To align Algorithm 2 with the notation of [8], set $G(L) = \frac{1}{2}\|L\|_F^2$ and

$$F(z) = \begin{cases} 0 & \text{if } \|z\|_2 \leq \eta; \\ \infty & \text{otherwise,} \end{cases}$$

i.e., F is the convex indicator of the scaled ℓ_2 norm ball. The convex conjugate F^* can be immediately computed: $F^*(y) = \eta\|y\|_2$ (see e.g. [27]). Then steps 5 and 7 of Algorithm 2

correspond to the *proximal updates*:

$$L^{k+1} = \arg\min_{L} \left\{ G(L) + \frac{1}{2\gamma}\|L - (L^k - \gamma \mathcal{A}^T y^k)\|^2 \right\};$$
$$y^{k+1} = \arg\min_{y} \left\{ F^*(y) + \frac{1}{2\gamma}\|y - y^+\|^2 \right\},$$

where y^+ is defined as in step 6 of Algorithm 2. The reader can check that any fixed point $(\overline{L}, \overline{y})$ of this iteration solves the saddle point formulation (3), and hence the problem in step 5 of Algorithm 1. See [8] for a detailed convergence analysis and rate results.

B. Relaxation of the η-constraint

We found that running Algorithm 1 with $\eta_k = \eta$, the final target value, is too aggressive, especially in the context of the alternating scheme in L and R. In particular, in the early steps, when the values of L and R are far from the optimum, it is to our advantage to solve approximate subproblems. As the algorithm progresses, we tighten η_k, and eventually solve the problem of interest with $\eta_k = \eta$.

V. Numerical Experiments

We perform seismic data interpolation using the proposed primal-dual alternating scheme on a 5D synthetic data set from a realistic complex geological model. Our goal is to interpolate 3D geological models with very fine-scale features and complex sedimentary environments (where interpolation problems are very challenging for coarsely sampled data), comparing level-set and primal-dual methods for residual-constrained formulations.

The 5D seismic synthetic data set is generated using the Compass velocity model provided to us by the BG Group, which has 101×101 receivers spaced by 25m, and 40×40 sources spaced by 25 also with temporal sampling interval of 0.004s (Figure 1). It consists of two receiver and two source coordinates, with time as the fifth dimension. We use ocean-bottom nodes style acquisition, where we places the receivers at the ocean-bed and sources are placed at sea-surface. In all experiments, we initialize L and R using standard Gaussian random matrices. More sophisticated initialization schemes have been proposed ([17], [1]), but are expensive for large-scale seismic data since they require computing the largest r singular vectors.

We compare the practical performance (signal-to-noise ratio (SNR) in dB of the reconstructions) and computational time (in seconds) of the new primal-dual formulation to that of LR-BPDN [1]. For simplicity, we perform interpolation for a missing-sources scenario, though the

missing-receivers scenario is similar. During seismic data acquisition, the sampling operator P_Ω only acts along source and receiver coordinates and does not depend on the time axis. We work with the data in Fourier domain as follows. We perform the Fourier transform along the time-axis, extract common monochromatic slices, transform them using the rank-revealing matricization, i.e., $\mathcal{S}$, and then interpolate each monochromatic data matrix independently. Once interpolation is finished for each monochromatic data slice, we apply $\mathcal{S}^H$ followed by a single inverse Fourier transform along the frequency axis to get the final reconstructed seismic data volume in the time-domain. This approach avoids repeated application of transformation operator $\mathcal{S}$, Fourier and inverse Fourier transforms during the optimization process.

We use jittered sampling scheme [13] to remove 80% of the sources resulting in the minimum spatial sampling of 25 m and maximum spatial sampling of 225 m between two consecutive sources. The spatial sampling of final interpolated grid is 25 m. Since seismic data is band-limited because of the band-limited nature of the source wavelet, we perform the interpolation within the frequency bandwidth of seismic data, which ranges from 3-70 Hz. In practice, successful recovery means that an actionable seismic volume (preserving all the coherent energy including late arrivals) can be obtained by computational techniques from merely 20% of acquired data.

To choose the rank of the factors L and R for interpolation, we consider the 5 Hz and 70 Hz frequency slices, subsample columns, and perform reconstruction. The best r values (according to SNR value) for 5 and 70 Hz monocromatic slices were 30 and 100, respectively. Hence, we work with all of the monochromatic slices and adjust the rank linearly from 30 to 100 when moving from lower to higher frequencies. To select the target value of the data fitting constraint η, we use the value $0.03\|b\|_F$, where b is the observed data. As described in Section IV-B, we start Algorithm 1 with a relatively loose value of $\eta_0 = \|b\|_F$, and decrease geometrically until we arrive at the target value of $0.03\|b\|_F$.

Figures 1 and 2 show fully sampled and subsampled (80% missing sources) data from the 5D seismic volume. Figures 1 (c) and 2 (c) show a zoom section of common-receiver gather from the true and subsampled data. The reconstructed data volume and corresponding residual plots are shown in Figures 6 (a, b) and (c,d) using the proposed alternating primal-dual splitting method and the LR-BPDN based level-set approach. We boost the amplitudes of residual plots (Figures 6 (c,d)) by a factor of 100 to show that we are not loosing any coherent energy. Both of the methods are able to reconstruct the coherent energy of 5D seismic data volumes while preserving the late-arrival energy, which can be also seen from the residual plots. We also plot

the zoom sections (Figure 7) of reconstructed data and corresponding residual to strengthen our observations.

We compare the signal-to-noise ratio (SNR in dB) and computational times (in sec.) across the frequency spectrum as shown in Figures 8 (a, b). We can see that the alternating primal-dual method works better along the higher frequency regime whereas the level-set method works better along the lower-frequency regime. Computationally, the primal-dual method is faster by a factor of two. In addition, the computational time for the primal-dual scheme is remarkably consistent across the frequency spectrum, while LR-BPDN requires more time as frequency increases. We run both the methods on a system with two 10-core Intel E5-2690 v2 Ivy Bridge CPUs at 3.00GHz.

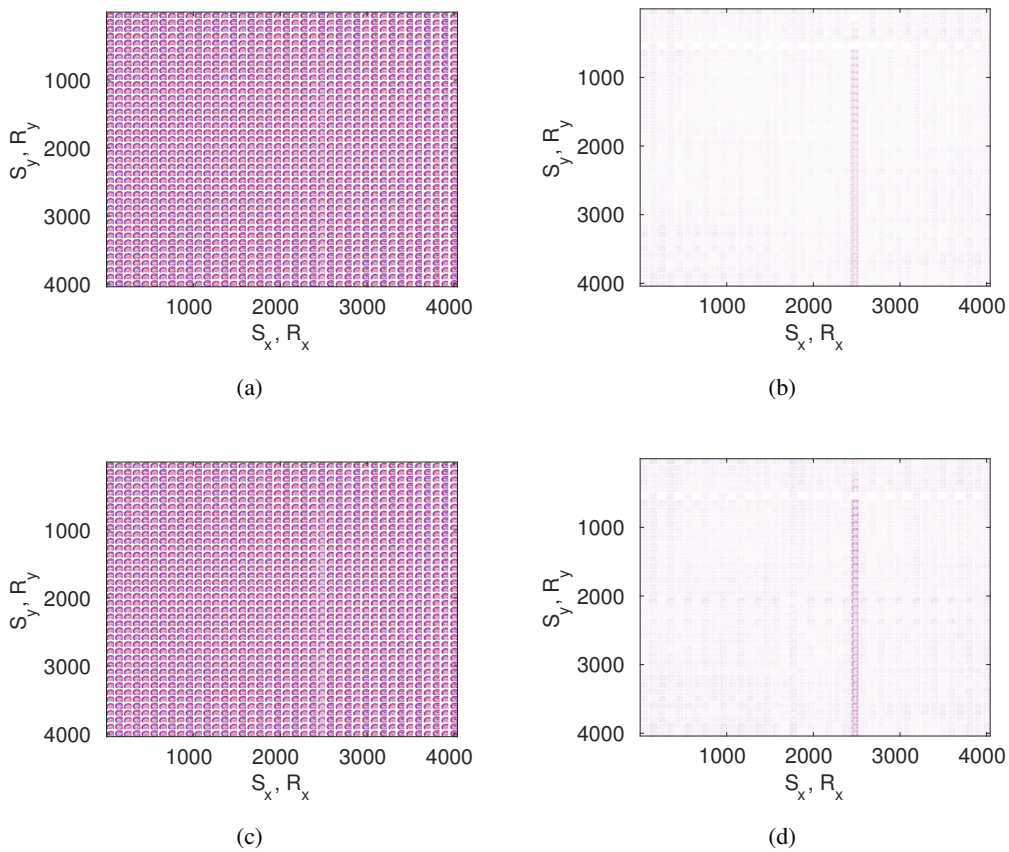


Fig. 5. Time-slices extracted at 1.0 s from interpolated data and its corresponding residual using (a,b) proposed methodology and (c,d) using level-set method from 5D seismic data volume for 80% missing sources scenario.

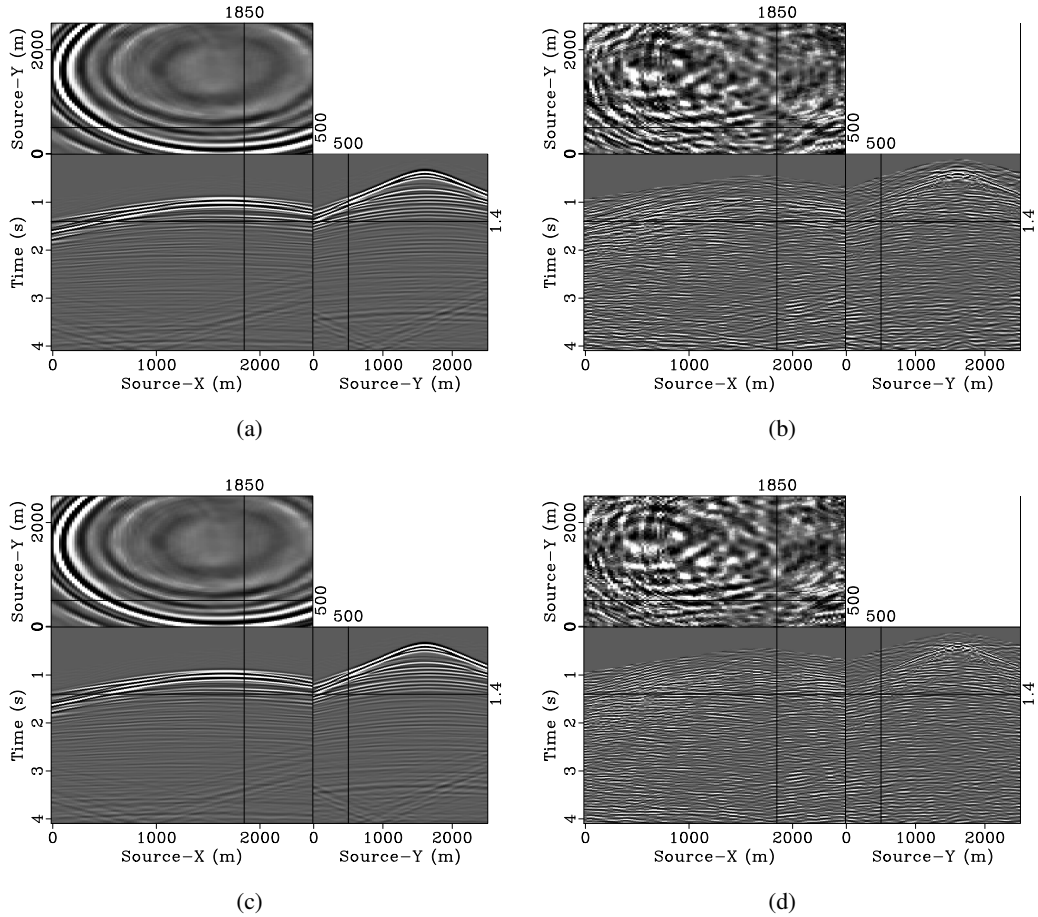


Fig. 6. A common-receiver gather from interpolated data and 100-times amplified residuals using (a,b) primal-dual splitting and (c,d) using level-set method for 80% missing sources scenario. We can clearly see that we recover most of the coherent energy using both the methods. We amplify the amplitude of residual by a factor of 100 to show that we are not loosing any coherent energy.

VI. Discussion and conclusions

Residual-constrained formulations are well suited for practical interpolation schemes, since they provide a simple way for practitioners to fit to a target data-fitting level. In this work, we propose a new approach for residual-constrained formulations, using a block coordinate alternating optimization scheme. As with other factorized schemes, in the current approach the full data volume is represented implicitly via the outer product $X = LR^H$, and achieves time and memory savings by working with the factors instead of the full volume X.

Even though the overall problem is nonconvex, each problem in L and R is a convex residual-constrained problem, and we solve it using a primal-dual splitting approach detailed in Algorithm 2. The approach requires only matrix-vector products with L, R and their adjoints. It is

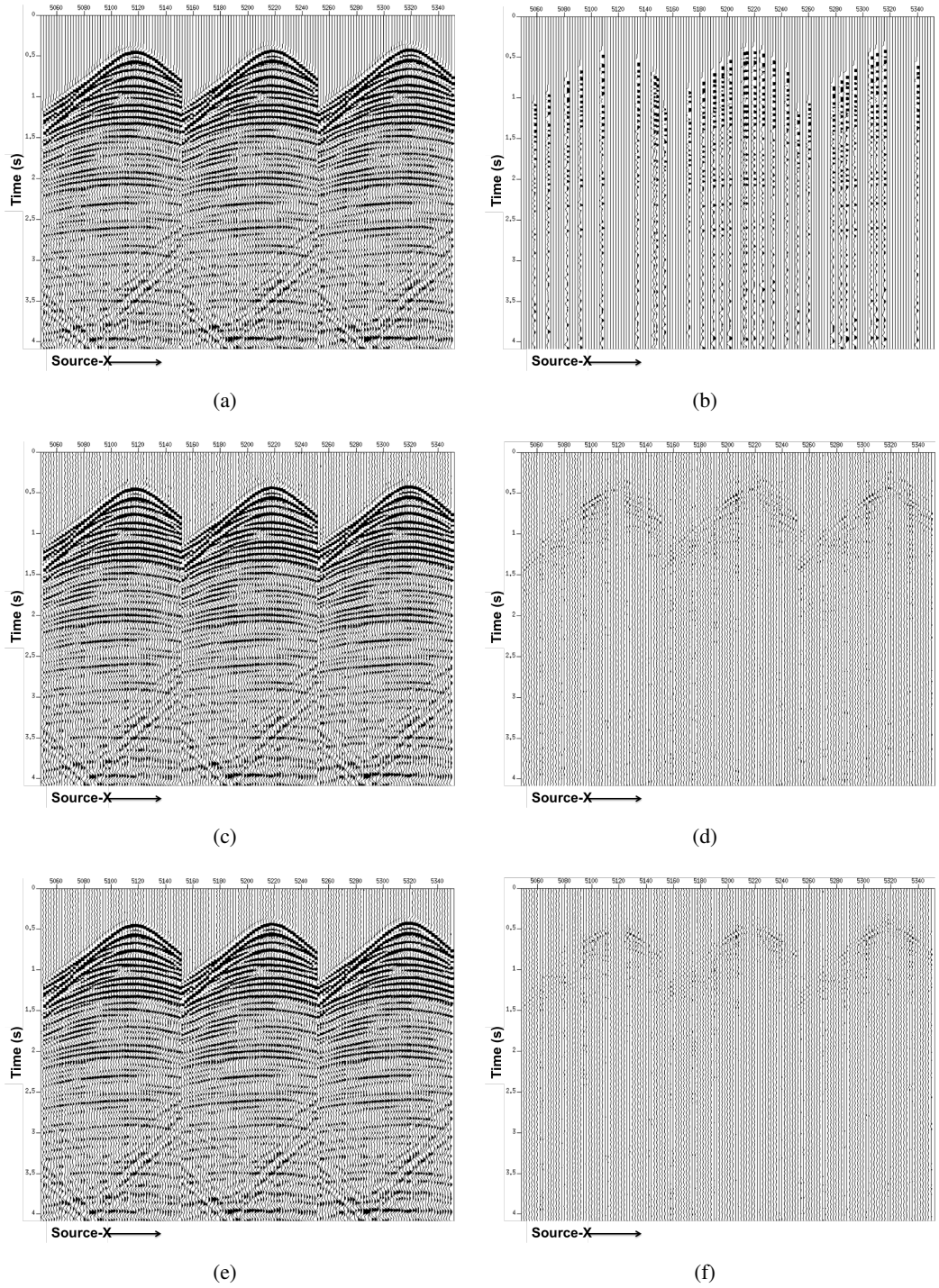
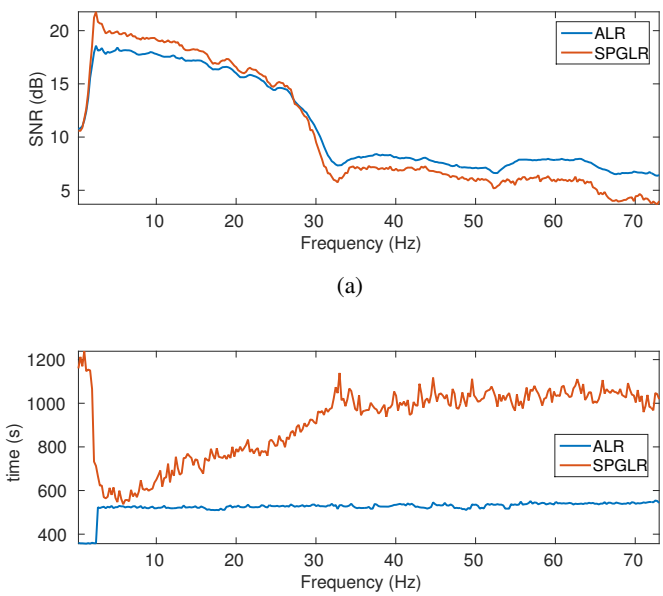


Fig. 7. Subsection of common-receiver gather (extracted from Figure 6). (a) Fully sampled ground truth. (b) 80% subsampled data. Interpolation results and corresponding residual using, (a,b) the proposed method, and (c,d) level-set method. Here, we keep the amplitude scale of reconstruction and residual same to show the efficacy of proposed formulation.



(a)

(b)

Fig. 8. Qualitative comparison of level-set method and the proposed primal-dual methodology over the frequency spectrum of 5D synthetic seismic data volume generated using BG Compass model for 80% missing sources. (a) signal-to-noise ratio (dB), (b) computational time (s).

simpler than the level-set approach, which requires variational computations of the value function. Experimental results on large-scale seismic interpolation show that the proposed approach can match the practical performance of level set methods, obtaining comparable results in half the time, and working consistently across the range of frequencies of interest.

An important consequence of this work is that it opens future directions in regularized data interpolation. One can for example consider constraints on the factors L, R themselves, or additional constraints on LR^H. When working with additional convex constraints, Algorithm 1 can be left completely unchanged, and Algorithm 2 can be adapted. The key point is that with L or R fixed, the subproblems remain convex. We leave these developments for future work.

ACKNOWLEDGEMENTS

We would like to thank the BG Group for permission to use the synthetic Compass velocity model. The authors wish to acknowledge the SENAI CIMATEC Supercomputing Center for Industrial Innovation, with support from BG Brazil and the Brazilian Authority for Oil, Gas and Biofuels (ANP), for the provision and operation of computational facilities and the commitment to invest in Research & Development. This work was financially supported in part by the

Natural Sciences and Engineering Research Council of Canada Collaborative Research and Development Grant DNOISE II (CDRP J 375142-08). This research was carried out as part of the SINBAD II project with the support of the member organizations of the SINBAD Consortium. Aleksandr Aravkin was partially supported by the Washington Research Foundation Data Science Professorship. Damek Davis was partially supported by the National Science Foundation under Award No. 1502405.

REFERENCES

[1] A. Aravkin, R. Kumar, H. Mansour, B. Recht, and F. J. Herrmann. Fast methods for denoising matrix completion formulations, with applications to robust seismic data interpolation. *SIAM Journal on Scientific Computing*, 36(5):S237–S266, 2014.

[2] A. Y. Aravkin, J. V. Burke, D. Drusvyatskiy, M. P. Friedlander, and S. Roy. Level-set methods for convex optimization. *arXiv preprint arXiv:1602.01506*, 2016.

[3] A. Y. Aravkin, J. V. Burke, and M. P. Friedlander. Variational properties of value functions. *SIAM Journal on optimization*, 23(3):1689–1717, 2013.

[4] E. Baysal, D. D. Kosloff, and J. W. Sherwood. Reverse time migration. *Geophysics*, 48(11):1514–1524, 1983.

[5] J.-F. Cai, E. J. Candès, and Z. Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 20(4):1956–1982, 2010.

[6] E. J. Candes and Y. Plan. Matrix completion with noise. *Proceedings of the IEEE*, 98(6):925–936, 2010.

[7] E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational mathematics*, 9(6):717–772, 2009.

[8] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.

[9] C. Da Silva and F. Herrmann. Hierarchical tucker tensor optimization-applications to 4d seismic data interpolation. In *75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013*, 2013.

[10] C. Da Silva and F. J. Herrmann. Optimization on the Hierarchical Tucker manifold - applications to tensor completion. *Linear Algebra and its Applications*, 481:131–173, 09 2015. (Linear Algebra and its Applications).

[11] M. Fazel. *Matrix rank minimization with applications*. PhD thesis, PhD thesis, Stanford University, 2002.

[12] T. Hastie, R. Mazumder, J. Lee, and R. Zadeh. Matrix completion and low-rank svd via fast alternating least squares. *arXiv preprint arXiv:1410.2596*, 2014.

[13] G. Hennenfent and F. J. Herrmann. Simply denoise: wavefield reconstruction via jittered undersampling. *Geophysics*, 73(3):V19–V28, 05 2008.

[14] F. J. Herrmann and G. Hennenfent. Non-parametric seismic data recovery with curvelet frames. *Geophysical Journal International*, 173(1):233–248, 2008.

[15] P. Jain, R. Meka, and I. S. Dhillon. Guaranteed rank minimization via singular value projection. In *Advances in Neural Information Processing Systems*, pages 937–945, 2010.

[16] P. Jain, P. Netrapalli, and S. Sanghavi. Low-rank matrix completion using alternating minimization. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 665–674. ACM, 2013.

[17] P. Jain, P. Netrapalli, and S. Sanghavi. Low-rank matrix completion using alternating minimization. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 665–674. ACM, 2013.

[18] N. Kreimer, A. Stanton, and M. D. Sacchi. Tensor completion based on nuclear norm minimization for 5d seismic data reconstruction. *Geophysics*, 78(6):V273–V284, 2013.

[19] R. Kumar, C. Da Silva, O. Akalin, A. Y. Aravkin, H. Mansour, B. Recht, and F. J. Herrmann. Efficient matrix completion for seismic data reconstruction. *Geophysics*, 80(5):V97–V114, 2015.

[20] T. T. Lin and F. J. Herrmann. Robust estimation of primaries by sparse inversion via one-norm minimization. *Geophysics*, 78(3):R133–R150, 05 2013.

[21] O. Lopez, R. Kumar, and F. J. Herrmann. Rank minimization via alternating optimization-seismic data interpolation. In *77th EAGE Conference and Exhibition 2015*, 2015.

[22] O. Lopez, R. Kumar, Ö. Yılmaz, and F. J. Herrmann. Off-the-grid low-rank matrix recovery and seismic data reconstruction. *IEEE Journal of Selected Topics in Signal Processing*, 10(4):658–671, 2016.

[23] H. Mansour, F. J. Herrmann, and Ö. Yılmaz. Improved wavefield reconstruction from randomized sampling via weighted one-norm minimization. *Geophysics*, 78(5):V193–V206, 2013.

[24] V. Oropeza and M. Sacchi. Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis. *Geophysics*, 76(3):V25–V32, 2011.

[25] R.-E. Plessix. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503, 2006.

[26] B. Recht, M. Fazel, and P. A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review*, 52(3):471–501, 2010.

[27] R. T. Rockafellar. *Convex Analysis*. Priceton Landmarks in Mathematics. Princeton University Press, 1970.

[28] N. Srebro, J. Rennie, and T. S. Jaakkola. Maximum-margin matrix factorization. In *Advances in neural information processing systems*, pages 1329–1336, 2004.

[29] S. Trickett, L. Burroughs, A. Milton, L. Walton, R. Dack, et al. Rank-reduction-based trace interpolation. In *2010 SEG Annual Meeting*. Society of Exploration Geophysicists, 2010.

[30] E. Van Den Berg and M. P. Friedlander. Probing the pareto frontier for basis pursuit solutions. *SIAM Journal on Scientific Computing*, 31(2):890–912, 2008.

[31] D. J. Verschuur, A. Berkhout, and C. Wapenaar. Adaptive surface-related multiple elimination. *Geophysics*, 57(9):1166–1177, 1992.

[32] J. Virieux and S. Operto. An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74(6):WCC1–WCC26, 2009.

[33] Y. Yang, J. Ma, and S. Osher. Seismic data reconstruction via matrix completion. *UCLA CAM Report*, pages 12–14, 2012.