Reconstruction of seismic wavefields via low-rank matrix factorization in the hierarchical-separable matrix representation

Rajiv Kumar, Hassan Mansour*, Aleksandr Y. Aravkin†, and Felix J. Herrmann**
University of British Columbia, Earth and Ocean Sciences, †IBM T.J. Watson Research Center

SUMMARY

Recent developments in matrix rank optimization have allowed for new computational approaches in the field of seismic data interpolation. In this paper, we propose an approach for seismic data interpolation which incorporates the Hierarchical Semi-Separable Structure (HSS) inside rank-regularized least-squares formulations for the missing-trace interpolation problem. The proposed approach is suitable for large scale problems, since it avoids SVD computations and uses a low-rank factorized formulation instead. We illustrate the advantages of the new HSS approach by interpolating a seismic line from the Gulf of Suez and compare the reconstruction with conventional rank minimization.

INTRODUCTION

Seismic data acquisition involves the recording of extremely large volumes of data, with the aim to inform seismic processing, such as surface related multiple estimation, estimation of primaries by sparse inversion, full waveform inversion and migration. In many situations, only a subset of the complete data is acquired due to physical, design, and/or budgetary constraints. Therefore, seismic missing-trace interpolation is one of the key pre-processing steps required to achieve high quality images of sub-surface structures. The trace interpolation problem is often solved using interpolation and denoising formulations that exploit data sparsity in some transform domain, for example using the Fourier (Sacchi et al. (1998)) and curvelet (Herrmann and Hennenfent (2008)) transforms. More recently, several works in the seismic literature have investigated data interpolation via rank minimization. The general idea is to exploit the low-rank structure of seismic data when it is organized in a matrix. *Low-rank structure* refers to the small number of nonzero singular values, or quickly decaying singular values. Oropeza and Sacchi (2011) identified that seismic temporal frequency slices organized into a block Hankel matrix, in ideal conditions, is a matrix of rank k, where k is the number of plane waves in the window of analysis. These authors showed that additive noise and missing samples increase the rank of the block Hankel matrix, and the authors presented an iterative algorithm that resembles seismic data reconstruction with the method of projection onto convex sets, where they use a low-rank approximation of the Hankel matrix via the randomized singular value decomposition (Liberty et al. (2007); Halko et al. (2011a)) to interpolate seismic temporal frequency slices. While this technique may be effective interpolating data with a limited number of distinct dips, the approach requires embedding the data into an even larger space where each dimension of size n is mapped to a matrix of size $n \times n$. To overcome the issue of the curse of dimensionality in higher dimensions, recent work by Silva and Herrmann (2013) has shown that high-dimensional seismic data can be recovered at low-frequencies by carrying out rank minimization using the Hierarchical Tucker format. They develop the rank minimization algorithm on the manifold of Hierarchical Tucker (HT) tensors for high-dimensional seismic data exhibiting low-rank structure represented as a tensor. The low-rank optimization formulation is solved on smooth manifolds using conjugate gradient methods for HT tensors.

In this paper, we extend the missing trace-interpolation method proposed by Aravkin et al. (2013) where interpolation is performed via a rank-regularized formulation. The key idea is that the monochromatic frequency slices of the fully sampled seismic-data matrix have low-rank structure when transformed into the midpoint-offset (m-h) domain, while the sub-sampled seismic-data matrix does not. That is, missing traces increase the rank of the resulting frequency slice in the m-h domain. The interpolation of seismic data is then achieved using a fast optimization approach that combines the (SVD-free) matrix factorization approach recently developed by Lee et al. (2010) with the Pareto curve approach proposed by Berg and Friedlander (2008). Kumar et al. (2013) extended this formulation to perform simultaneous seismic data interpolation and denoising using a robust rank-regularized formulation.

One of the main requirements of exploiting rank-minimization approaches is that the target data set should exhibit a low-rank structure. Seismic frequency slices exhibit low-rank structure in the m-h domain at the low-frequencies, but not at the high-frequencies. This behaviour is due to the increase in oscillations as we move from low to high-frequency slices in the m-h domain, even though the energy remains focused around the diagonal. Therefore, interpolation via rank minimization in the high-frequency range requires extended formulations that incorporate low-rank structure. While Engquist and Ying (2010) propose to overcome this difficulty by proposing directional low-rank approximations for Green's functions of the acoustic wave equation in 3D, we rely instead on the Hierarchically Semi-Separable matrix representation (HSS) method proposed by Chandrasekaran et al. (2006) to represent frequency slices. The key idea in the HSS representation is that certain full-rank matrices, e.g., matrices that are diagonally dominant with off-diagonals that decay away from the diagonal, can be represented by a collection of low-rank sub-matrices. Jumah and Herrmann (2012) showed that HSS representations can be used to reduce the storage and computational cost for the estimation of primaries by sparse inversions. They combined the HSS representation with the randomized SVD proposed by Halko et al. (2011b) to accelerate matrix vector multiplications that are required for sparse inversion.

In this paper, we present a methodology to improve trace interpolation by combining HSS with factorization-based rank-regularized optimization formulations (Lee et al., 2010). We first show that we can represent monochromatic frequency slices of seismic data in terms of low-rank sub-matrices in the m-h domain. Next, we demonstrate the efficacy of the new HSS based method by interpolating frequency slices from a seismic line from the Gulf of Suez and compare the reconstruction with the conventional rank minimization formulation.

METHODOLOGY

The missing trace interpolation problem can be perceived as a matrix completion problem. Let X_0 be a matrix in $\mathbb{C}^{n\times m}$ and let $\mathscr{A}$ be a linear measurement operator that maps from $\mathbb{C}^{n\times m} \rightarrow \mathbb{C}^p$ with $p \ll n\times m$. When the target matrix X_0 is low-rank and given a set of measurements $b = \mathscr{A}(X_0)$, the matrix completion problem can be solved by finding the minimum rank solution X that satisfies the system of equations $\mathscr{A}(X) = b$. However, the rank minimization problem is NP-hard and is computationally intractable, in general. Recht et al. (2010) showed that under certain general conditions on the operator $\mathscr{A}$, the solution to the rank minimization problem can be found by solving the following nuclear-norm minimization problem:

$$\min_X ||X||_* \quad \text{subject to} \quad \|\mathscr{A}(X) - b\|_2 \leq \varepsilon, \qquad (\text{BPDN}_\sigma)$$

where $\|X\|_* = \|\sigma\|_1$, and σ is the vector of singular values.

The frequency slices of seismic data organized in a seismic line generally do not exhibit a low-rank structure in the source-receiver domain since strong wavefronts extend diagonally across the source-receiver plane. However, transforming the data into the midpoint-offset domain results in a vertical alignment of the wavefronts, thereby reducing the rank of the frequency slice matrix. In the lower frequency slices, this vertical alignment can be accurately approximated by a low-rank representation. On the other hand, higher frequency slices include a variety of wave oscillations that increase the rank of the high-frequency slices. To illustrate this phenomenon, we plot the decay of the singular values for low and high-frequency slices from the Gulf of Suez seismic line depicted in Figure 1. The figure shows that the singular value decay rate is slower for the high-frequency slice than for the low-frequency slice. However, it is possible to find accurate low-rank approximations of submatrices of the high-frequency slices by partitioning the data into a Hierarchical Semi-Separable (HSS) structure. The HSS structure first partitions a matrix into diagonal and off-diagonal submatrices. The same partitioning structure is then applied recursively to the diagonal submatrices only. To illustrate the HSS partitioning we consider a 2D monochromatic high-frequency data matrix at 60Hz in the s-r domain. We show the first level of partitioning in Figure 2(a) and the second level partitioning in Figure 2(b) in their corresponding source-receiver domains. For the same data, we display each (first-level) diagonal and off-diagonal sub-block in Figure 3(a) and the decay of the singular values for each (first-level) sub-block in Figure 3(b) in their corresponding source-receiver and midpoint-offset domains. We can clearly see that the off-diagonal sub-matrices have low-rank structure, while the diagonal sub-matrices have higher rank. Further partitioning of the diagonal sub-blocks allows us to find better low-rank approximations as shown in Figure 2(b).

When the seismic data have missing traces in the source-receiver domain, the rank of the subsampled matrix is not necessarily reduced compared to the fully sampled data. In fact, the vertical alignment of the missing traces in the source-receiver domain often results in a data matrix with a lower rank than the fully sampled data. However, rotating the data by 45 degrees into the midpoint-offset domain results in a lower rank representation of the fully sampled data and missing traces induce a higher rank for the subsampled data. To illustrate this behaviour, we plot the decay of the singular values of fully sampled and subsampled monochromatic frequency slice in the source-receiver and midpoint-offset domains for 12Hz and 60Hz frequency slices, respectively in Figure 4. Notice how subsampling does not noticeably change the decay of the singular values in the source-receiver domain as expected, but significantly slows down the decay rate in the midpoint-offset domain. Therefore, low-frequency slices can be reconstructed by solving the nuclear-norm minimization problem in the midpoint-offset domain, while high-frequency slices can be reconstructed by first performing an HSS partitioning and then solving the nuclear-norm minimization problem on each partition in its respective midpoint-offset domain.

In this paper, we avoid the direct approach to nuclear-norm minimization that involves computing singular value decompositions (SVD) of the matrices and follow instead the approach proposed by Aravkin et al. (2013). The approach efficiently solves (BPDN_σ) using an extension of the SPGℓ_1 solver (Berg and Friedlander, 2008) using theoretical results from Aravkin et al. (2012). To speed up the computation, we adopt a recent factorization-based approach to nuclear-norm minimization introduced by (Rennie and Srebro (2005); Lee et al. (2010); Recht and Ré (2011)). The factorization approach parametrizes the matrix $X \in \mathbb{C}^{n\times m}$ as the product of two low-rank factors $L \in \mathbb{C}^{n\times k}$ and $R \in \mathbb{C}^{m\times k}$, such that, $X = LR^H$, where the superscript H indicates the Hermitian transpose. The optimization is done using the factors L and R instead of X, thereby significantly reducing the size of the decision variable from nm to $k(n+m)$ when $k \ll m,n$. In (Rennie and Srebro, 2005), it was shown that the nuclear-norm obeys the relationship

$$\|X\|_* = \|LR^T\|_* \leq \frac{1}{2}\left\|\begin{bmatrix} L \\ R \end{bmatrix}\right\|_F^2 =: \Phi(L,R), \qquad (1)$$

where $\|\cdot\|_F^2$ is Frobenius norm of the matrix (sum of the squared entires). Consequently, the LASSO subproblem can be replaced by

$$\min_{L,R} \|\mathscr{A}(LR^T) - b)\|_2 \quad \text{subject to} \quad \Phi(L,R) \leq \tau\,, \qquad (2)$$

where the projection onto $\Phi(L,R) \leq \tau$ is easily achieved by multiplying each factor L and R by the scalar $2\tau/\Phi(L,R)$. By equation (1) for each HSS sub-matrix in the m-h domain, we are guaranteed that $\|LR^T\|_* \leq \tau$ for any solution of equation (2). The extension of Berg and Friedlander (2008) discussed in Aravkin et al. (2013) allows us to embed this idea into the Pareto curve approach, traversing an approximate Pareto curve for (BPDN_σ) while working with factors L,R. Once the optimization problem is solved, we transform each sub-matrix back from the m-h to the s-r domain, where we concatenate all the sub-matrices to get the interpolated monochromatic frequency data matrices. One of the advantages of HSS representation is that it works with the recursive partitioning of a matrix, and sub-matrices can be solved in parallel, speeding up the optimization formulation.

NUMERICAL EXPERIMENTS

We implement the proposed formulation on a seismic line from the Gulf of Suez with $N_s = 352$ sources, $N_r = 352$ receivers, $N_t = 1024$ time samples and a sampling interval of 0.004s. Most of the energy of the seismic line is concentrated in the 12-60Hz frequency band. In order to perform interpolation, we apply a sub-sampling mask that randomly removes 50% of the shots. In this example, we perform three levels of HSS partitioning on each frequency slice in the s-r domain. We then perform the interpolation on each sub-block separately by first transforming each sub-block into the m-h domain, and then solving the nuclear-norm minimization formulation (BPDN$_\sigma$) on each sub-block for a fixed rank of 5 (for L, R in equation (2)). We use 300 iterations of SPGℓ_1 for all frequency slices. For comparison, we also performed the interpolation with the conventional rank minimization formulation. We can see in Figure 5 that interpolation results significantly improve when we incorporate the HSS structure. We also compare the advantage of including the HSS structure by using one, two, and three levels of partitioning. As shown in Figure 6, just incorporating the first-level of HSS partitioning significantly improves the signal-to-noise ratio (SNR). We also observe that HSS can also be used for low-frequency data matrices but there is no significant difference in recovery for the low-frequencies data matrices as shown in Figure 6. This behaviour is due to the fact that low-frequency data matrices exhibit low-rank structure, therefore we can find accurate low-rank approximations with conventional rank minimization formulation.

DISCUSSION

We have presented a new method for seismic missing trace interpolation, where we incorporate the HSS struture into low-rank trace interpolation by (factorized) nuclear-norm minimization. We combine the Pareto curve approach for optimizing (BPDN$_\sigma$) formulations with the SVD-free matrix factorization methods to solve the nuclear-norm optimization formulation. The resulting formulation can be solved in a straightforward way by using the extended SPGℓ_1 algorithm (Aravkin et al., 2013).

Incorporating HSS into this scheme is very promising, since we can attack high-rank structures using the partitioning, and in addition gain a computational advantage since optimization can be done in parallel on each block. The experimental results demonstrate the potential benefit of this methodology.

ACKNOWLEDGEMENTS

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE II (375142-08).This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BGP, BP, Chevron, ConocoPhillips, Petrobras, ION GXT, CGGVeritas, Woodside, PGS, Total SA, and WesternGeco.

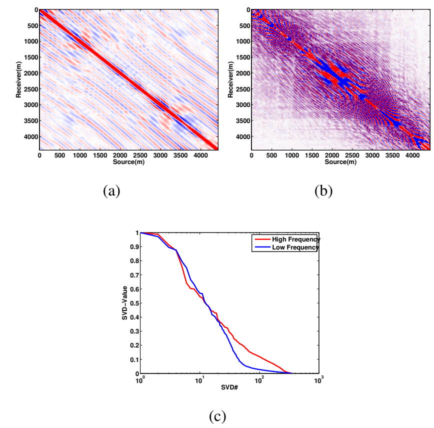

Figure 1: (a) Low-frequency data matrices at 12 Hz and (b) High-frequency data matrices at 60Hz in s-r domain. (c) Singular-value decay of low and high-frequency data matrices in s-r domain.

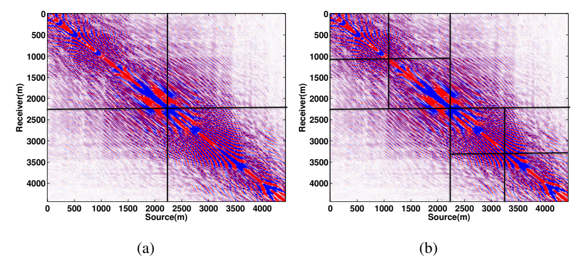

Figure 2: HSS partitioning of a high-frequency data matrices at 60 Hz in s-r domain. (a) First level and (b) Second level of HSS partitioning.

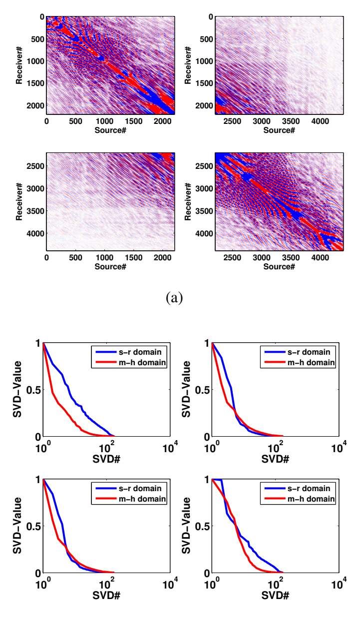

Figure 3: Sub-blocks after first level of HSS partitioning of a high-frequency data matrices at 60 Hz in s-r domain. (a) Diagonal and off-diagonal sub-blocks. (b) Decay of singular values for each sub-block in s-r and m-h domain.

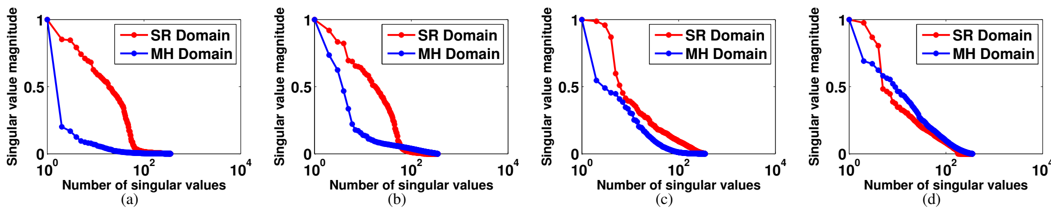

Figure 4: Singular value decay of fully sampled (a) low-frequency slice at 12 Hz and (c) high-frequency slice at 60 Hz in s-r and m-h domains. Singular value decay of 50% subsampled (b) low-frequency slice at 12 Hz and (d) high-frequency slice at 60 Hz in s-r and m-h domains.

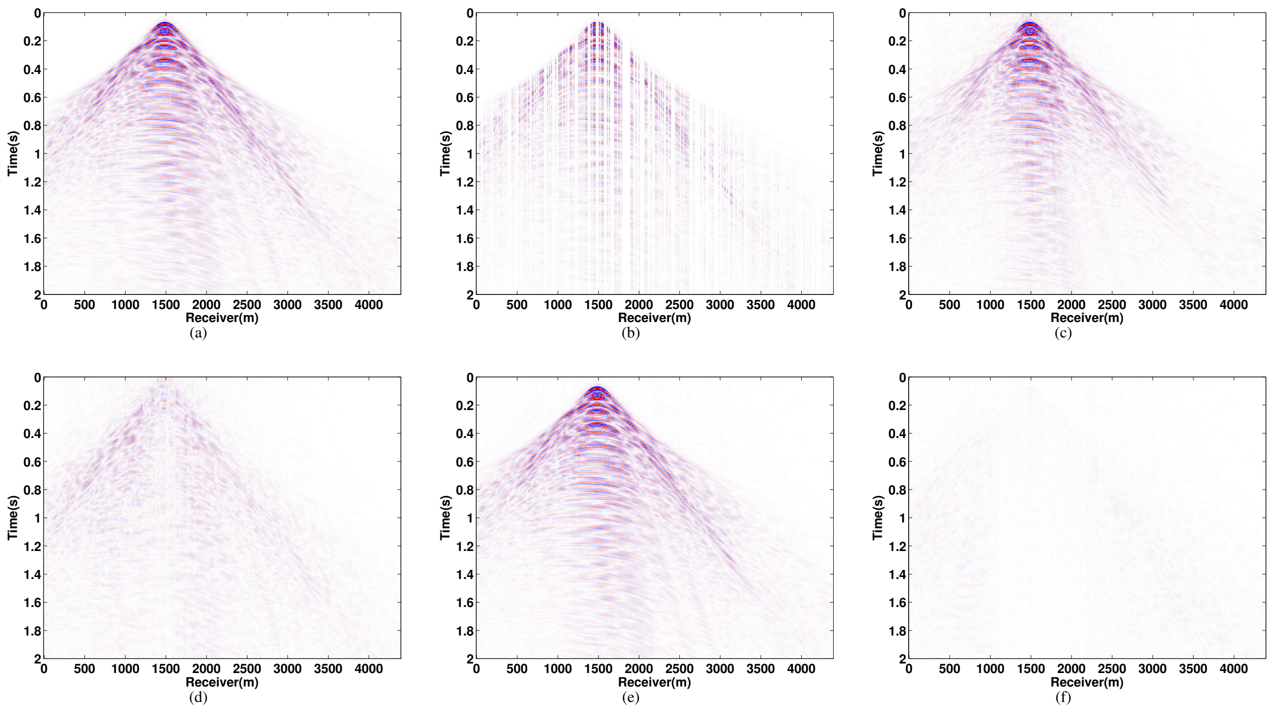

Figure 5: Seismic data interpolation with and without HSS. (a,b) Fully sampled and subsampled data with subsampling ratio of 50%. (c,d) Recovery and residual without HSS with a SNR of 11 dB. (e,f) Recovery and residual using HSS structure with a SNR of 19 dB.

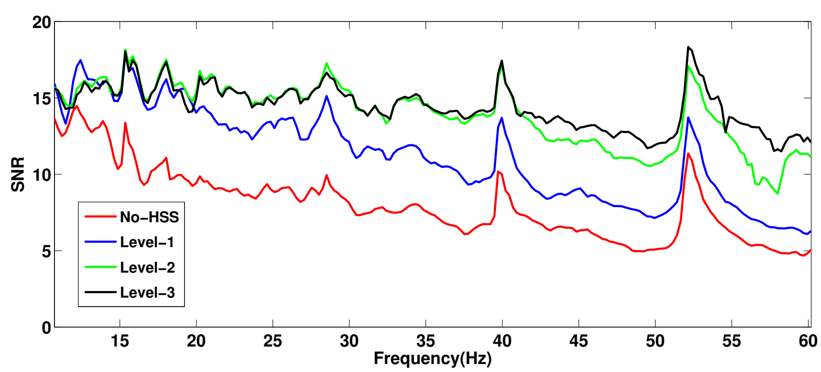

Figure 6: Qualitative measurement of seismic data interpolation for 12-60 Hz frequency band. We can see the significant improvements in the SNR by incorporating the HSS structure.

REFERENCES

Aravkin, A., J. Burke, and M. Friedlander, 2012, Variational properties of value functions: submitted to SIAM J. Opt., arXiv:1211.3724.

Aravkin, A. Y., R. Kumar, H.Mansour, B. Recht, and F. Herrmann, 2013, An svd-free pareto curve approach to rank minimization: arXiv:1302.4886.

Berg, E. v., and M. P. Friedlander, 2008, Probing the pareto frontier for basis pursuit solutions: SIAM Journal on Scientific Computing, **31**, 890–912.

Chandrasekaran, S., P. Dewilde, M. Gu, W. Lyons, and T. Pals, 2006, A fast solver for hss representations via sparse matrices: SIAM J. Matrix Analysis Applications, **29**, 67–81.

Engquist, B., and L. Ying, 2010, Fast directional algorithms for the helmholtz kernel: J. Comput. Appl. Math., **234**, 1851–1859.

Halko, N., P.-G. Martinsson, and J. A. Tropp, 2011a, Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions: SIAM Review, **53**, 217–288.

———, 2011b, Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions: SIAM Review, **53**, 217–288.

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

Jumah, B., and F. J. Herrmann, 2012, Dimensionality-reduced estimation of primaries by sparse inversion.

Kumar, R., A. Y. Aravkin, H.Mansour, B. Recht, and F. Herrmann, 2013, Seismic data interpolation and denoising using svd-free low-rank matrix factorization: Presented at the , EAGE.

Lee, J., B. Recht, R. Salakhutdinov, N. Srebro, and J. Tropp, 2010, Practical large-scale optimization for max-norm regularization: Presented at the Advances in Neural Information Processing Systems, 2010.

Liberty, E., F. Woolfe, P.-G. Martinsson, V. Rokhlin, and M. Tygert, 2007, Randomized algorithms for the low-rank approximation of matrices: **104**, 20167–20172.

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

Recht, B., M. Fazel, and P. Parrilo, 2010, Guaranteed minimum rank solutions to linear matrix equations via nuclear norm minimization.: SIAM Review, **52**, 471–501.

Recht, B., and C. Ré, 2011, Parallel stochastic gradient algorithms for large-scale matrix completion: Presented at the Optimization Online.

Rennie, J. D. M., and N. Srebro, 2005, Fast maximum margin matrix factorization for collaborative prediction: Proceedings of the 22nd international conference on Machine learning, ACM, 713–719.

Sacchi, M., T. Ulrych, and C. Walker, 1998, Interpolation and extrapolation using a high-resolution discrete fourier transform: Signal Processing, IEEE Transactions on, **46**, 31 –38.

Silva, C. D., and F. J. Herrmann, 2013, Hierarchical tucker tensor optimization - applications to 4d seismic data interpolation: Presented at the , EAGE.