

High-frequency wavefield recovery with weighted matrix factorizations

Yijun Zhang¹

Shashin Sharan²

Felix J. Herrmann^{1,2}

¹ Department of Electrical & Computer Engineering, Georgia Institute of Technology

² Department of Earth & Atmospheric Sciences, Georgia Institute of Technology

SUMMARY

Acquired seismic data is normally not the fully sampled data we would like to work with since traces are missing due to physical constraints or budget limitations. Rank minimization is an effective way to recovering the missing trace data. Unfortunately, this technique's performance may deteriorate at higher frequency because high-frequency data can not necessarily be captured accurately by low-rank matrix factorizations albeit remedies exist such as hierarchical semi-separable matrices. As a result, recovered data often suffers from low signal to noise ratio S/Rs at the higher frequencies. To deal with this situation, we propose a weighted recovery method that improves the performance at the high frequencies by recursively using information from matrix factorizations at neighboring lower frequencies. Essentially, we include prior information from previously reconstructed frequency slices during the wavefield reconstruction. We apply our method to data collected from the Gulf of Suez, which shows that our method performs well compared to the traditional method without weighting.

INTRODUCTION

Seismic data acquisition is one of the key steps in initial phase of oil & gas exploration. Due to operational complexity and operational costs, acquired seismic data is usually not fully sampled, a prerequisite to subsequent steps such as multiple removal and migration all of which require densely sampled data.

Wavefield recovery is an important tool to solve the problem of poor sampling. In the last decade, wavefield recovery methods based on sparsity promotion in different transform domains, such as the Radon (Bardan, 1987), wavelet (Villasenor et al., 1996), and curvelet (Herrmann et al., 2007; Herrmann and Hennefent, 2008) domain have been developed. Although these methods are valuable in terms of the quality of recovered data, they are relatively complex and computationally expensive. Fortunately, matrix completion methods (Kumar et al., 2015) based on low-rank matrix factorizations are relatively simple and computationally cheaper. The latter use the property that fully-sampled frequency slices permit accurate low-rank representations when organized in midpoint/offset. In Kumar et al. (2015); Oropeza and Sacchi (2011), authors exploit the fact that presence of noise or missing traces increases the rank of these frequency slices. We use this property to recover frequency slices via factored rank minimization (Kumar et al., 2014). While this matrix factorization method performs well at the low to mid frequencies, it struggles to recover high-frequency data, which need higher ranks to be accurately represented.

Recent work by Aravkin et al. (2014); Eftekhari et al. (2018)

has shown that reliable prior information on the row and column subspaces of the underlying low rank matrix can be used to further improve wavefield recovery via matrix completion. For seismic data, we have access to this information when there is a strong similarity between adjacent frequency slices. In that case, the row and column subspaces can serve as prior information. This principle was first demonstrated by Aravkin et al. (2014) and we extend this line of research by recursively invoking prior as we work our way from the relatively low frequencies to the high frequencies where conventional matrix completion methods typically perform poorly.

We organize our paper as follows. First, we discuss wavefield recovery via matrix completion. Next, we discuss how to incorporate prior information on the row and column subspaces from neighboring lower frequencies in our matrix completion framework. We conclude demonstrating our approach on a field data example from the Gulf of Suez and show its better performance compared to conventional matrix completion especially at the higher frequencies.

METHODOLOGY

Low-rank matrix factorization

In Kumar et al. (2015) and Aravkin et al. (2014), authors exploit low rank of fully sampled seismic data by solving for each frequency problem of the type

$$\underset{\mathbf{X}_i}{\text{minimize}} \quad \|\mathbf{X}_i\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X}_i) - \mathbf{b}_i\|_2 \leq \varepsilon. \quad (1)$$

In this expression, the matrices \mathbf{X}_i for $i = 1 \dots n_f$ with n_f the number of angular frequencies represent fully sampled monochromatic frequency slices in the midpoint/offset domain, \mathcal{A} is the sampling operator collecting the data into a vector, and \mathbf{b}_i represents the observed data at the i^{th} frequency. For each frequency, we recover the fully sampled data by minimizing the nuclear norm $\|\cdot\|_*$ on each \mathbf{X}_i subject we fit the data within ε . The nuclear norm itself is defined as the sum of the singular values. We solve Equation 1 for all the frequencies to obtain our recovered data $\mathbf{X} \in \mathbb{C}^{n_f \times n_m \times n_h}$, where n_m is the number of midpoints and n_h the number of offsets. As reported by Kumar et al. (2015), randomized sampling increases the rank of 2D seismic data in midpoint offset domain, which is a favorable condition for matrix completion.

To avoid computationally expensive singular-value decompositions (SVD) while solving 1, we employ a low-rank matrix factorization approach. For this purpose, we factor the matrices (for notational simplicity we drop the subscript i) $\mathbf{X} \in \mathbb{C}^{n_m \times n_h}$ in Problem 1 into the low-rank factors $\mathbf{L} \in \mathbb{C}^{n_m \times r}$ and $\mathbf{R} \in \mathbb{C}^{n_h \times r}$ both of rank r . To avoid expensive SVDs, we follow Rennie and Srebro (2005) and replace the nuclear norm in Problem 1

High-frequency wavefield recovery

by

$$\underset{\mathbf{L}, \mathbf{R}}{\text{minimize}} \quad \frac{1}{2} \left\| \begin{bmatrix} \mathbf{L} \\ \mathbf{R} \end{bmatrix} \right\|_F^2 \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{L}\mathbf{R}^H) - \mathbf{b}\|_F \leq \varepsilon, \quad (2)$$

where H is the Hermitian transpose and $\|\cdot\|_F$ the Frobenius norm (2-norm of the vectorized matrix). Following Kumar et al. (2015) and Aravkin et al. (2014), we solve Problem 2 with spectral-projected gradients (Van Den Berg and Friedlander, 2008).

As we mentioned earlier, the performance of low-rank factorization methods degrade with increasing frequency reflected in increasing poor signal to noise ratios (SNRs) (Kumar et al., 2015). To improve the recovered data quality at higher frequencies, we include weighted matrix completion (Aravkin et al., 2014; Eftekhari et al., 2018).

Weighted low-rank matrix factorization

The key of our methodology is that we approximate fully sampled data in low-rank factored form. When using SVDs, this factored form reads

$$\mathbf{X} \approx \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H, \quad (3)$$

where $\mathbf{U} \in \mathbb{C}^{n_m \times r}$ and $\mathbf{V} \in \mathbb{C}^{n_h \times r}$ are column and row subspaces of \mathbf{X} , respectively. $\mathbf{\Sigma} \in \mathbb{C}^{r \times r}$ is a diagonal matrix containing the largest r singular values of the matrix \mathbf{X} .

As shown by Aravkin et al. (2014); Eftekhari et al. (2018), information on these subspaces can be used to rewrite Problem 1 into its weighted form—i.e., we have

$$\underset{\mathbf{X}}{\text{minimize}} \quad \|\mathbf{Q}\mathbf{X}\mathbf{W}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X}) - \mathbf{b}\|_2 \leq \varepsilon, \quad (4)$$

where

$$\mathbf{Q} = w_1 \mathbf{U}\mathbf{U}^H + \mathbf{U}^\perp \mathbf{U}^{\perp H} \quad (5)$$

and

$$\mathbf{W} = w_2 \mathbf{V}\mathbf{V}^H + \mathbf{V}^\perp \mathbf{V}^{\perp H} \quad (6)$$

are projection matrices on subspaces spanned by \mathbf{U} , \mathbf{V} and their orthogonal complements \mathbf{U}^\perp , \mathbf{V}^\perp . The scalars $w_1 \in (0, 1]$ and $w_2 \in (0, 1]$ are weights that depend on the confidence we have in the priors—i.e., how close the matrix \mathbf{X} is to the actual \mathbf{X}_i we are dealing with at frequency i . Small values for the weights w_1 and w_2 mean that we have confidence in the prior (the matrix \mathbf{X} is close). When $w_1 \uparrow 1$ and $w_2 \uparrow 1$, solving Problem 4 becomes equivalent to solving the original Problem 1.

As before, we can rewrite Problem 4 into a weighted low-rank factored form (Aravkin et al., 2014):

$$\underset{\mathbf{L}, \mathbf{R}}{\text{minimize}} \quad \frac{1}{2} \left\| \begin{bmatrix} \mathbf{Q}\mathbf{L} \\ \mathbf{W}\mathbf{R} \end{bmatrix} \right\|_F^2 \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{L}\mathbf{R}^H) - \mathbf{b}\|_2 \leq \varepsilon. \quad (7)$$

The question now is which subspaces to use for the columns and rows. Because frequency slices have information in common from frequency to frequency, we follow Aravkin et al. (2014); Eftekhari et al. (2018) and use the \mathbf{U} and \mathbf{V} from the previous lower frequency.

While Aravkin et al. (2014) somewhat successfully applied this approach for a single frequency slice, these authors never justified this approach and neither did they apply the weighting recursively working from the low to the higher frequencies. For this purpose, we quantify similarities (in the form of angles, see Eftekhari et al. (2018)) between subsequent frequency slices for the Gulf of Suez data. This will allow us to predict the performance of our method.

Quantifying similarity

Similarity between subsequent frequency slices depends on the largest principal angles (Eftekhari et al., 2018) between column subspaces and between row subspaces of subsequent frequency slices. Smaller angles correspond to more similarity between subspaces of subsequent frequency slices and vice versa. Therefore, we can choose smaller weights w_1 and w_2 when the angles are smaller. Smaller weights correspond to larger penalties (Eftekhari et al., 2018) on matrices that have subspaces orthogonal to \mathbf{U} and \mathbf{V} in 4. When weights are small, we have more confidence in \mathbf{U} and \mathbf{V} and less confidence in their orthogonal counterparts. In Figure 1, we show angles between column subspaces (Figure 1b) and row subspaces (Figure 1a) for subsequent frequency slices of the Gulf of Suez data. We observe an overall decreasing trend in angles with increasing frequencies for both row and column subspaces. This trend indicates increasing similarity between subsequent frequency slices with increasing frequency. This trend is consistent with high-frequency approximate behavior of wavefields—i.e., as the frequency increases solutions become more and more like the high-frequency solution, and this gives us a handle how to choose the weights as the frequency increases.

Figure 1 shows that as the frequency increase, the largest angles between the subspaces of neighboring frequencies decreases. Smaller the angle, more similar are subspaces. This angle test have demonstrated that the weighted matrix completion will perform better in high frequency band because of smaller angle in comparison to its lower frequency counterpart.

NUMERICAL EXPERIMENTS

We demonstrate effectiveness of recursive weighted matrix completion method over the conventional matrix completion method and the weighted method just using previous frequency slice (named as single pair weighted method) in terms of data reconstruction quality. In single pair weighted method, we reconstruct previous frequency slice using conventional matrix completion. We use real 2D seismic data with number of sources, $N_s = 354$, and number of receivers, $N_r = 354$ acquired in Gulf of Suez for this comparison. Total number of time samples for this data set is $N_t = 1024$ with sampling interval of 0.004 s. Most of the energy of the seismic line is concentrated in 20 Hz to 70 Hz frequency band (Figure 3b). To get the subsampled data (Figures 2b & 4b), we remove 75% of sources using a jittered subsampling mask. Jittered subsampling method not only breaks the inherent properties such as low rank of fully sampled seismic data but also controls the maximum gap size of the incomplete data (Herrmann and Hennenfent, 2008). We show the results and comparisons in both frequency domain

High-frequency wavefield recovery

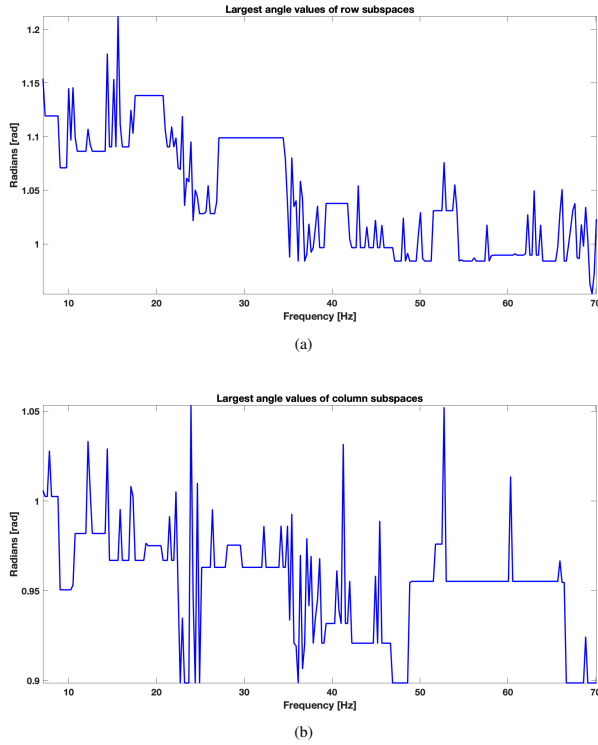


Figure 1: Largest angle between (a) row and (b) column subspaces for subsequent frequency slices of Gulf of Suez data

and time domain. For every frequency slice, we perform 150 iterations of spectral projected gradient algorithm for all these methods. Figure 2 shows results on a frequency slice at 30Hz. Recovery with the unweighted method gives the result with S/R of 11.43 dB. Whereas recovery with the single pair weighted method gives a higher S/R of 15.20 dB, our recursive weighted method gives the highest S/R of 18.48 dB, 7.05 dB improvement in S/R over the unweighted method. Figure 2d, Figure 2f and Figure 2h show differences between these three different methods and ground truth data. The reconstruction using recursive prior knowledge gives the least residual in Figure 2h among these three methods. The residual of reconstruction without using any prior knowledge is Figure 2d and the residual of single pair weighted reconstruction is Figure 2f.

Figure 3a shows recursive weighted method's performance (red color plot) over range of frequencies in terms of signal to noise of completion. Recursive weighted recovery clearly outperforms the conventional recovery without weight (black color plot in Figure 3a) and the single pair weighted recovery (blue color plot in Figure 3a) in the frequency range which contains most of the energy (Figure 3b).

To further compare these three methods for time domain data, we apply them on all frequency slices and transfer results back to time domain. Figure 4 shows results and differences for one common receiver gather extracted from complete time domain data. With the conventional recovery we get S/R of 6.87 dB, whereas with single pair weighted recovery we get S/R of 7.85 dB and with recursive weighted recovery we get S/R of

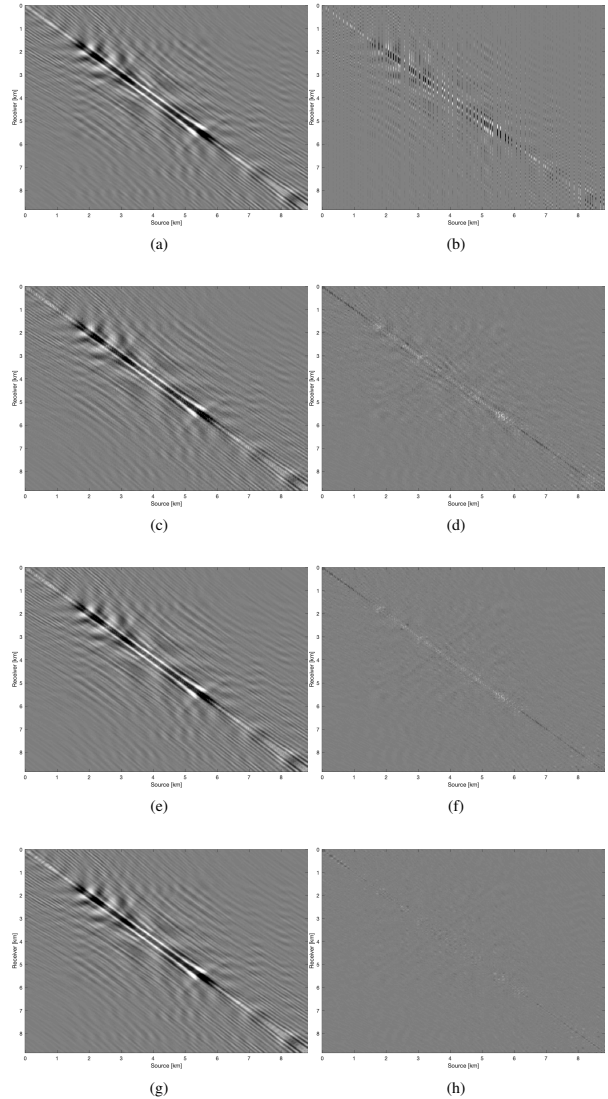


Figure 2: Missing trace recovery for a frequency slice at 30Hz in source-receiver domain. (a) Ground truth, (b) 75% subsampled seismic data with jittered subsampling. (c) and (d) represent recovery ($S/R = 11.43$ dB) using conventional method and its difference w.r.t. the ground truth respectively. (e) and (f) represent recovery ($S/R = 15.20$ dB) using single pair weighted method and its difference w.r.t. the ground truth respectively. (g) and (h) represent recovery ($S/R = 18.48$ dB) using recursive weighted method and its difference w.r.t. the ground truth respectively.

11.63 dB. With recursive weighted recovery we get almost 5 dB improvement for complete time domain data over conventional recovery. It is also obvious to see the advantage of the recursive weighted method from three differences in Figure 4d, Figure 4f and Figure 4h. The residual is significantly reduced in Figure 4h in contrast to Figure 4d and Figure 4f.

High-frequency wavefield recovery

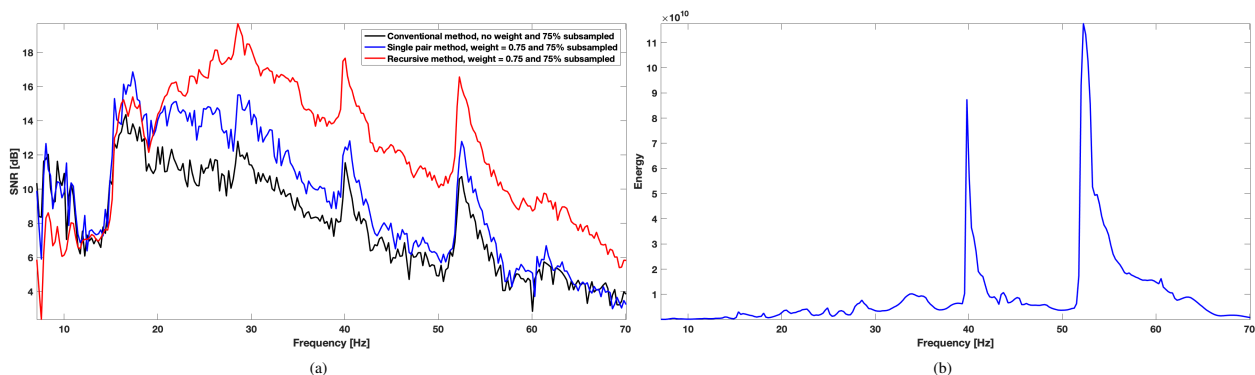


Figure 3: (a) S/R vs frequency of recovery using recursive weighted method (Red color), single pair weighted method (Blue color) and conventional method (Black color). (b) Plot of energy of frequency slices vs frequency.

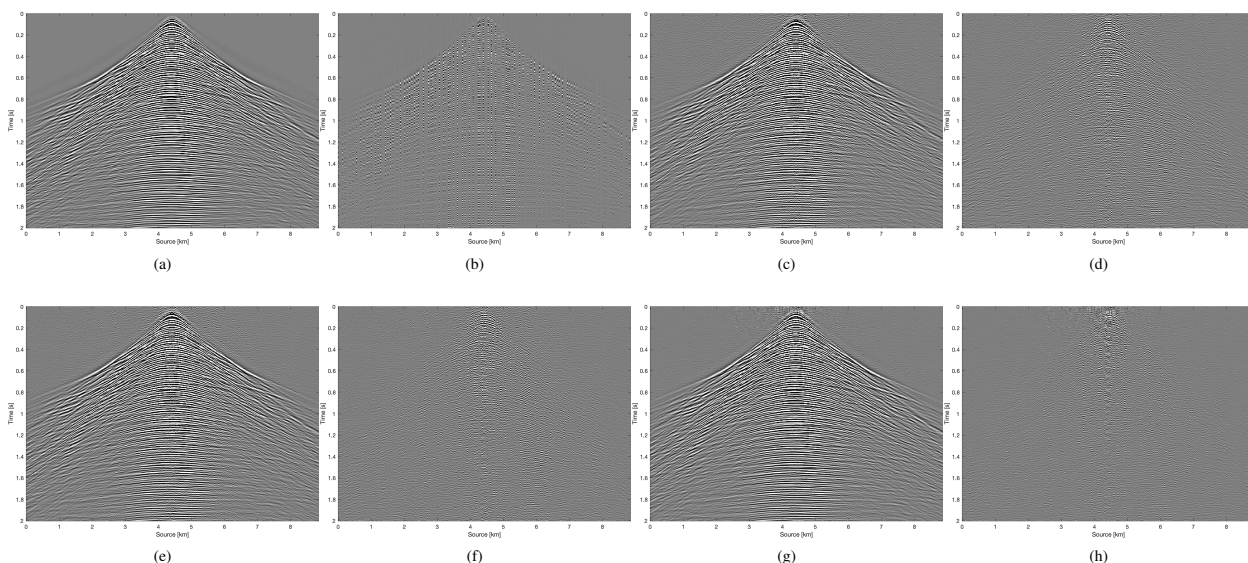


Figure 4: Missing trace recovery of time domain data. (a) Ground truth. (b) 75% subsampled seismic data with jittered subsampling. (c) and (d) represent recovery ($S/R = 6.87$ dB) using conventional method and its difference w.r.t. the ground truth respectively. (e) and (f) represent recovery ($S/R = 7.85$ dB) using single pair weighted method and its difference w.r.t. the ground truth respectively. (g) and (h) represent recovery ($S/R = 11.63$ dB) using recursive weighted method and its difference w.r.t. the ground truth respectively.

CONCLUSIONS

In this work, we have proposed recursive weighted matrix completion to improve data reconstruction quality especially at high frequencies where the conventional matrix completion method performs poorly. In contrast to conventional low-rank matrix factorization without weighting or with non-recurrent pairwise weighting, our recursively weighted method performs better at the high frequencies, especially at frequencies where the data has the most energy. We also demonstrated the effectiveness of our recursive recovery on real data. Future work will be to extend the application of recursive weighted matrix completion to realistic size 3D seismic data reconstruction and also to simultaneous source acquisition.

ACKNOWLEDGEMENT

The first and second authors have equal contribution to this work. This research was carried out as part of the SINBAD II project with the support of the member organizations of the SINBAD Consortium.

REFERENCES

- Aravkin, A., R. Kumar, H. Mansour, B. Recht, and F. J. Herrmann, 2014, Fast methods for denoising matrix completion formulations, with applications to robust seismic data interpolation: *SIAM Journal on Scientific Computing*, **36**, S237–S266.
- Bardan, V., 1987, Trace interpolation in seismic data processing: *Geophysical Prospecting*, **35**, 343–358.
- Eftekhari, A., D. Yang, and M. B. Wakin, 2018, Weighted matrix completion and recovery with prior subspace information: *IEEE Transactions on Information Theory*, **64**, 4044–4071.
- Herrmann, F. J., and G. Hennenfent, 2008, Non-parametric seismic data recovery with curvelet frames: *Geophysical Journal International*, **173**, 233–248.
- Herrmann, F. J., D. Wang, G. Hennenfent, and P. P. Moghaddam, 2007, Curvelet-based seismic data processing: A multiscale and nonlinear approach: *Geophysics*, **73**, A1–A5.
- Kumar, R., A. Y. Aravkin, E. Esser, H. Mansour, and F. J. Herrmann, 2014, Svd-free low-rank matrix factorization-wavefield reconstruction via jittered subsampling and reciprocity: 76th EAGE Conference and Exhibition 2014, 1–5.
- Kumar, R., C. Da Silva, O. Akalin, A. Y. Aravkin, H. Mansour, B. Recht, and F. J. Herrmann, 2015, Efficient matrix completion for seismic data reconstruction: *Geophysics*, **80**, V97–V114.
- Oropeza, V., and M. Sacchi, 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: *Geophysics*, **76**, V25–V32.
- Rennie, J. D., and N. Srebro, 2005, Fast maximum margin matrix factorization for collaborative prediction: *Proceedings of the 22nd international conference on Machine learning*, ACM, 713–719.
- Van Den Berg, E., and M. P. Friedlander, 2008, Probing the pareto frontier for basis pursuit solutions: *SIAM Journal on Scientific Computing*, **31**, 890–912.
- Villasenor, J. D., R. Ergas, and P. Donoho, 1996, Seismic data compression using high-dimensional wavelet transforms: *Proceedings of Data Compression Conference-DCC'96*, IEEE, 396–405.