Data denoising and interpolation using synthesis and analysis sparse regularization

Lucas Almeida[*1], *Michael Wakin*[2] *& Paul Sava*[1]
[1]*Center for Wave Phenomena, Colorado School of Mines*
[2]*Department of Electrical Engineering and Computer Science, Colorado School of Mines*

SUMMARY

Missing trace reconstruction is a challenge in seismic processing due to incomplete and irregular acquisition. Noise is a concern, due to the many sources of noise that occur during seismic acquisition. Most of the recent research on denoising and interpolation focuses on transform domain approaches using L_1 norm minimization. A specific kind of constraint, called the synthesis approach, is widely used in geophysical problems. The analysis approach, which can be considered as the synthesis' dual problem, is an alternative for sparsity-constrained inversion. Although less popular than the synthesis solution, the analysis approach is more effective in several problems, such as denoising of natural images. We compare and contrast the analysis and synthesis approaches as sparsity constraints for the missing trace reconstruction and denoising problems. We show that the analysis approach yields more accurate results than the synthesis approach, which makes it a viable approach for sparsity-constrained inversion for noise related geophysical problems.

INTRODUCTION

Seismic acquisition ideally seeks to sample densely and regularly in every spatial direction, with the intent of obtaining signals that adequately represent data observed at the surface. Good sampling has important consequences for many applications, such as reverse-time migration and multiple removal. However, acquisition costs and field obstacles can easily make the acquisition both irregular and sparse. The purpose of missing trace reconstruction is to fill the data gaps or to resample the data as accurately as possible.

Transform domains for missing trace estimation have been proposed by several authors, (Hennenfent and Herrmann, 2008; Naghizadeh and Sacchi, 2010). This approach assumes that the analyzed wavefield is sparse in some domain and involves the minimization of a convex function with an L_1 norm, which is known to promote sparsity. This approach is also used to denoise complex data corrupted by incoherent noise (Herrmann et al., 2007; Zhu et al., 2015). Hennenfent and Herrmann (2008) suggest that both problems are very similar when the considered transform domain is Fourier related, as missing trace reconstruction can be regarded as a denoise problem in the Fourier domain.

The transform domain approach is usually related to the compressive sensing problem (Candès et al., 2006). In this setting, one tries to find the sparsest representation that describes the signal to be estimated in a transform domain. This formulation is known as the **synthesis approach** and it has been widely used for several geophysical problems, such as full-waveform inversion (Li et al., 2012), deblending (Wason et al., 2011) and salt body detection (Ramirez et al., 2016).

The **analysis approach** provides a different way of solving the L_1 minimization problem by estimating the signal in the ambient domain using a set of forward transforms. Although less popular than the synthesis approach, the analysis approach and its geometry has been studied theoretically from the perspective of compressive sensing (Candes et al., 2011) and the cosparse model (Nam et al., 2013). In geophysics, the analysis approach has been used sparingly to solve problems such as denoising with data-driven frames (Chen et al., 2016) and multiple removal (Yang and Fomel, 2015).

We compare the analysis and synthesis approaches for missing trace reconstruction and for denoising, and illustrate both techniques with a complex shot gather from the Sigsbee 2A model, for different spatial sampling schemes.

THEORY

The underconstrained inverse problem can be formulated mathematically as

$$\mathbf{y} = \mathbf{\Phi x}, \tag{1}$$

where $\mathbf{\Phi}$ is a linear operator that maps a high dimensional vector $\mathbf{x} \in \Re^n$ to a low dimensional vector $\mathbf{y} \in \Re^m$. A general minimization problem that estimates $\mathbf{x}$ is (Cai et al., 2012)

$$\min_{\mathbf{u}} \quad \|\mathbf{\Phi\Theta}^*\mathbf{u} - \mathbf{y}\|_2 + \kappa\|(\mathbf{I} - \mathbf{\Theta\Theta}^*)\mathbf{u}\|_2 + \gamma\|\mathbf{u}\|_1, \tag{2}$$

where κ and γ are regularization parameters. In equation 2, $\mathbf{x}$ is assumed to be sparse with respect to a transform $\mathbf{\Theta}$, and the estimated signal in the ambient domain is recovered by $\mathbf{x} = \mathbf{\Theta}^*\mathbf{u}$.

Two special cases of equation 2 are of particular interest in the literature. If $\kappa = 0$, then the equation simplifies to

$$\min_{\mathbf{u}} \quad \gamma\|\mathbf{u}\|_1 + \|\mathbf{y} - \mathbf{\Phi\Theta}^*\mathbf{u}\|_2. \tag{3}$$

This equation defines the synthesis approach, because the estimated signal is synthesized using $\mathbf{x} = \mathbf{\Theta}^*\mathbf{u}$. When $\kappa \longrightarrow \infty$, the second term of equation 2 is zero, which gives rise to the analysis approach

$$\min_{\mathbf{x}} \quad \gamma\|\mathbf{\Theta x}\|_1 + \|\mathbf{y} - \mathbf{\Phi x}\|_2. \tag{4}$$

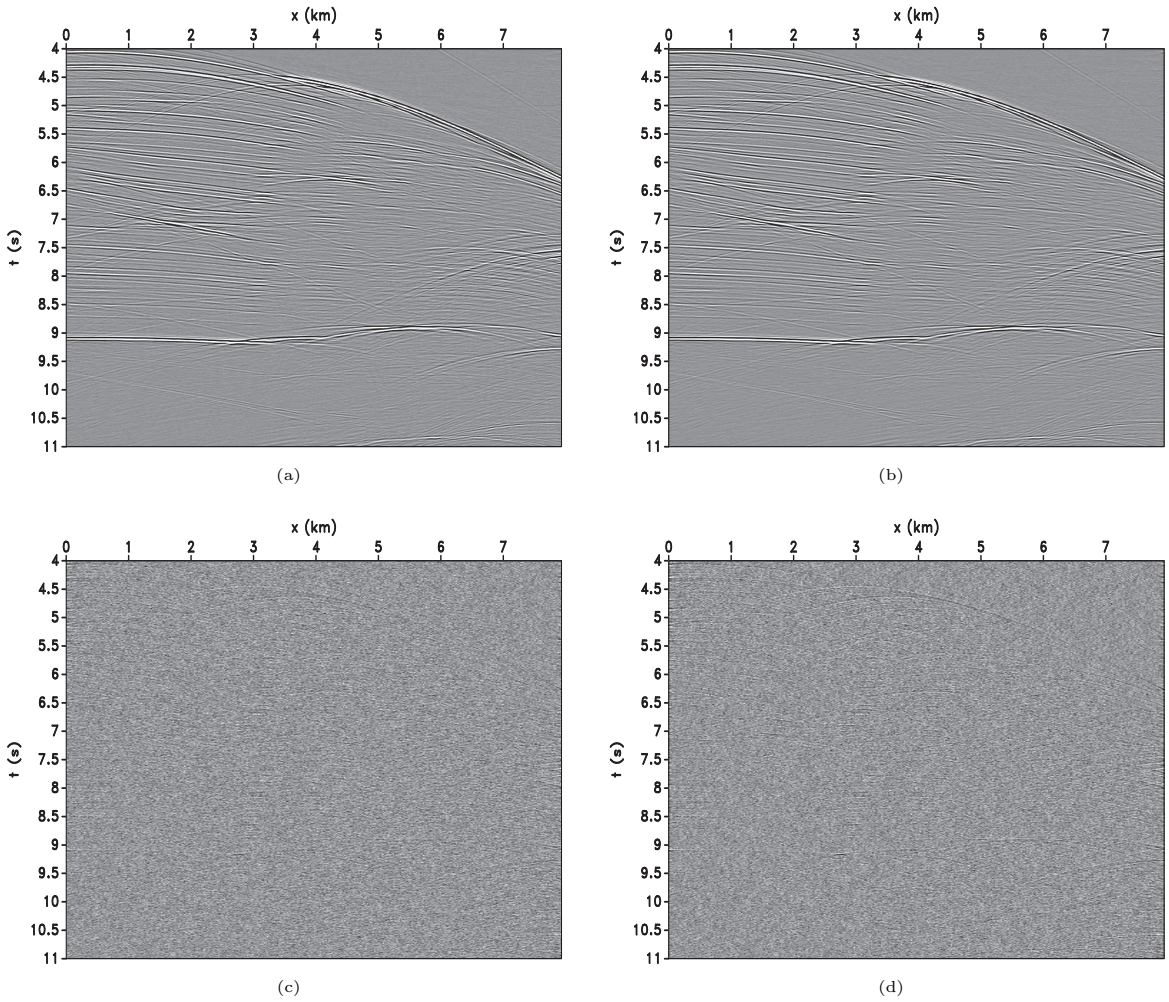


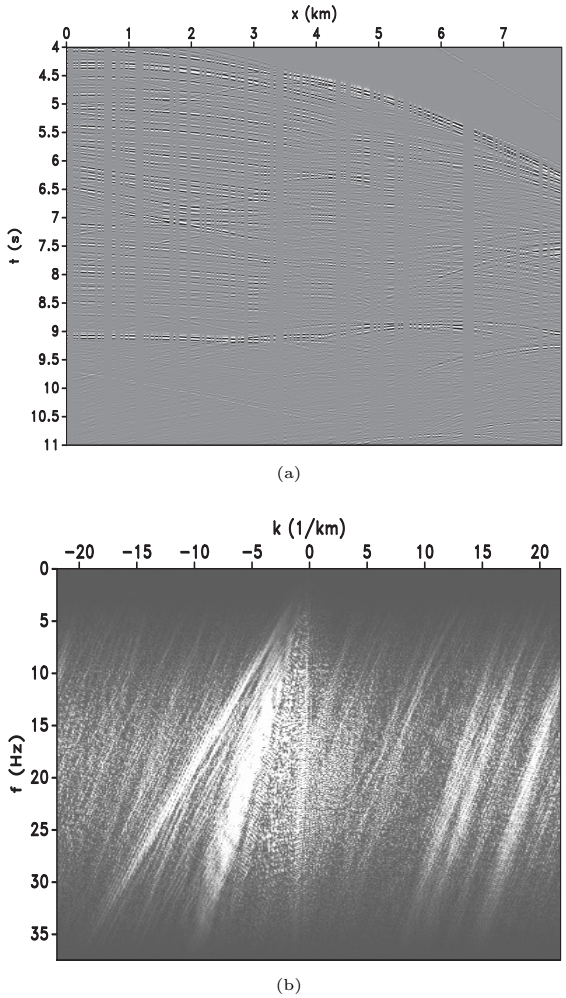
Figure 1: Denoised shot gathers: (a) Synthesis and (b) analysis solutions with SNR of 17.46 dB and 18.22 dB, respectively. (c), (d) Corresponding difference plots for panels (a) and (b), respectively.

If $\mathbf{\Theta}$ is an isometry, both formulations yield the same solution. However, in the general case, this is not true because $\mathbf{\Theta}\mathbf{\Theta}^* \neq \mathbf{I}$. One of the main differences between the approaches comes from the fact that the solutions are estimated in different domains: while the synthesis approach seeks a sequence $\mathbf{u}$ in the transform domain, the analysis approach seeks to find a dataset $\mathbf{x}$ in the ambient domain. In particular, the solution of equation 3 is closely related to the soft-threshold operator (Daubechies et al., 2004), which implies that it must be the sparsest $\mathbf{u}$ that represents a given $\mathbf{x}$ in the transform domain. On the other hand, the solution of equation 4 is dominated by the second term in equation 2 as $\kappa \longrightarrow \infty$. Thus, it is related to the orthogonal projector $\mathbf{I} - \mathbf{\Theta}\mathbf{\Theta}^*$, which implies that this formulation seeks the sparsest $\mathbf{u}$ that belongs to the null space of this orthogonal projector. Thus, in general, the solution of the synthesis approach is much sparser than that of the analysis approach.

More sparsity can be an advantage for the synthesis approach in terms of descriptive power, because its solution needs fewer elements to describe a signal in the transform domain when compared to that of the analysis approach. However, this can be a drawback when we consider erroneous estimation: every coefficient of the synthesis solution carries more significance than those of the analysis solution. Thus, the misestimation of the magnitude or support of coefficients of the synthesis solution, for example, in the presence of noise, might yield a solution that is very different from the desired result (Elad et al., 2007). This motivates us to study the performance of both approaches when geophysical problems involving noise are considered.

EXAMPLES

We evaluate the performance of equations 3 and 4 when applied to the missing trace reconstruction and denois-



(a)

(b)

Figure 2: (a) Shot gather after random undersampling, and (b) corresponding f-k spectrum.

ing problems. We specify $\mathbf{\Phi}$ as a restriction matrix that spatially undersamples the data by 50% for the former and as the identity matrix for the latter, while $\mathbf{\Theta}$ is taken as the curvelet transform (Candes et al., 2006). We use a complex shot gather from the Sigsbee 2A model, which features a wide range of amplitudes and conflicting dips. To solve both approaches in similar grounds, we use the NESTA solver (Becker et al., 2011) for all the experiments.

For the denoising problem, we contaminate the considered synthetic shot gather with random (incoherent) Gaussian noise. Figure 1 shows the solution and corresponding difference plot for the synthesis and analysis approaches. We can see that both approaches do well in removing the noise present in the data, although the difference plots also show some damage to the seismic reflections that should stay untouched. This is common in transform domain approaches, even when data-driven transforms are used (Chen et al., 2016). However, the SNR of the estimated solutions indicates that the analysis approach obtains a better solution than the synthesis approach for removing incoherent noise.

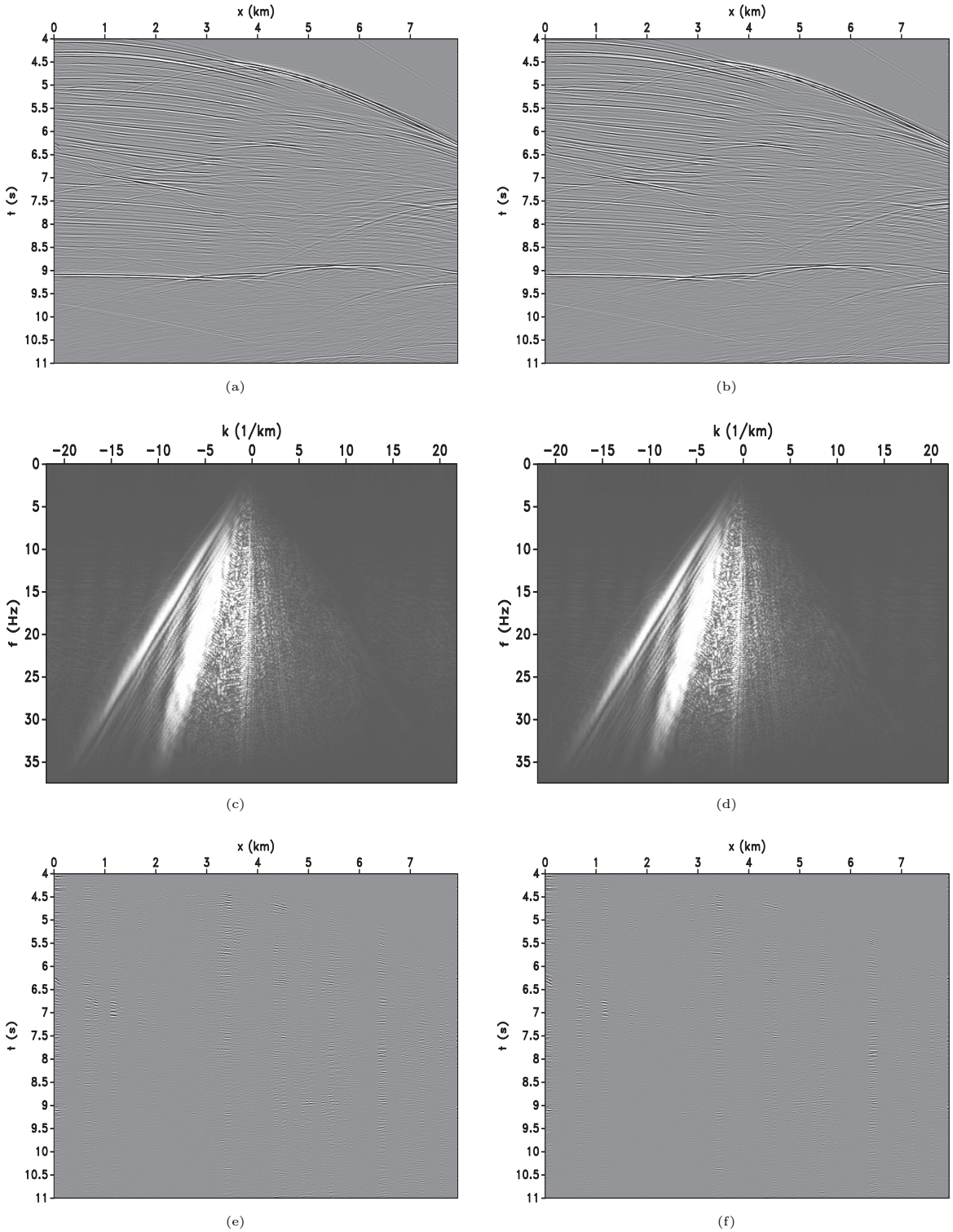
For the missing trace reconstruction problem, we evaluate the results for three undersampling schemes: uniform, random and jittered (Hennenfent and Herrmann, 2008), but in the following we illustrate only the random sampling results, for simplicity. Figures 2a and 2b show the randomly undersampled data and the corresponding f-k spectra. Note that, although the f-k spectrum features no coherent aliases, the resulting spectral leakage from the random undersampling assumes a somewhat coherent pattern, i.e., it is not completely incoherent noise as in the prior application. Because we use the curvelet transform as the transform domain, we can consider the interpolation of the missing traces as the denoising of the noisy f-k spectrum. Figure 3 shows the solutions for the synthesis and analysis approaches along with the corresponding f-k spectra and difference plots for the random undersampling scheme. The SNR for the uniform and jittered undersampling experiments are: 13.29 dB and 16.26 dB for the synthesis solutions; 14.83 dB and 17.93 dB for the analysis solutions. The SNR and difference plots confirm that the analysis approach is superior to the synthesis approach for every undersampling scheme.

CONCLUSIONS

The state-of-the-art solutions for the denoising and missing trace reconstruction problems are based on transform domain approaches. In the context of sparsity-promoting inversion, we identify two key techniques: the analysis and the synthesis approaches. We compare these methods on two noise related problems. For the denoising case, our examples show that the analysis approach is superior to the synthesis approach, obtaining a clean image with less damage to the signal and thus attaining a higher SNR than that of the synthesis approach. For the missing trace reconstruction problem, our examples show that the analysis approach consistently finds better solutions than the synthesis counterpart for any of undersampling scheme. Our analysis indicates that the analysis approach is a strong candidate for solving inverse problems in geophysical applications involving noise when a sparsity-promoting constraint is efficient and it should be considered along with the synthesis approach.

ACKNOWLEDGMENTS

We thank the sponsors of the Center for Wave Phenomena, whose support made this research possible. We also thank the authors of CurveLab and NESTA for making codes available. The reproducible numeric examples in this paper use the Madagascar open-source software package (Fomel et al., 2013) freely available from www.ahay.org.



(a) (b) (c) (d) (e) (f)

Figure 3: Random undersampling: (a), (b) Synthesis and analysis solutions. SNR of 16.24 dB and 17.87 dB, respectively. (c), (d) Corresponding f-k spectra. (e), (f) Corresponding difference plots scaled by 10.

REFERENCES

Becker, S., J. Bobin, and E. J. Candès, 2011, Nesta: a fast and accurate first-order method for sparse recovery: SIAM Journal on Imaging Sciences, **4**, 1–39.

Cai, J.-F., B. Dong, S. Osher, and Z. Shen, 2012, Image restoration: Total variation, wavelet frames, and beyond: Journal of the American Mathematical Society, **25**, 1033–1089.

Candes, E., L. Demanet, D. Donoho, and L. Ying, 2006, Fast discrete curvelet transforms: Multiscale Modeling & Simulation, **5**, 861–899.

Candes, E. J., Y. C. Eldar, D. Needell, and P. Randall, 2011, Compressed sensing with coherent and redundant dictionaries: Applied and Computational Harmonic Analysis, **31**, 59–73.

Candès, E. J., J. Romberg, and T. Tao, 2006, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information: Information Theory, IEEE Transactions on, **52**, 489–509.

Chen, Y., J. Ma, and S. Fomel, 2016, Double sparsity dictionary for seismic noise attenuation: Geophysics, **81**, V103–V116.

Daubechies, I., M. Defrise, and C. De Mol, 2004, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint: Communications on pure and applied mathematics, **57**, 1413–1457.

Elad, M., P. Milanfar, and R. Rubinstein, 2007, Analysis versus synthesis in signal priors: Inverse problems, **23**, 947.

Fomel, S., P. Sava, I. Vlad, Y. Liu, and V. Bashkardin, 2013, Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments: Journal of Open Research Software, **1**.

Hennenfent, G., and F. J. Herrmann, 2008, Simply denoise: wavefield reconstruction via jittered undersampling: Geophysics, **73**, V19–V28.

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.

Li, X., A. Y. Aravkin, T. van Leeuwen, and F. J. Herrmann, 2012, Fast randomized full-waveform inversion with compressive sensing: Geophysics, **77**, A13–A17.

Naghizadeh, M., and M. D. Sacchi, 2010, Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data: Geophysics, **75**, WB189–WB202.

Nam, S., M. E. Davies, M. Elad, and R. Gribonval, 2013, The cosparse analysis model and algorithms: Applied and Computational Harmonic Analysis, **34**, 30–56.

Ramirez, C., G. Larrazabal, and G. Gonzalez, 2016, Salt body detection from seismic data via sparse representation: Geophysical Prospecting, **64**, 335–347.

Wason, H., F. J. Herrmann, T. T. Lin, et al., 2011, Sparsity-promoting recovery from simultaneous data: A compressive sensing approach: Presented at the 2011 SEG Annual Meeting, Society of Exploration Geophysicists.

Yang, P., and S. Fomel, 2015, Seislet-based morphological component analysis using scale-dependent exponential shrinkage: Journal of Applied Geophysics, **118**, 66–74.

Zhu, L., E. Liu, and J. H. McClellan, 2015, Seismic data denoising through multiscale and sparsity-promoting dictionary learning: Geophysics, **80**, WD45–WD57.