Reweighted thresholding and orthogonal projections for simultaneous source separation

Satyakee Sen, Zhaojun Liu, James Sheng, and Bin Wang, TGS*

Summary

The key to success for simultaneous source separation is the ability to formulate an appropriate sparse inversion problem so that nontrivial solutions to a highly under-determined system can be found. An important issue with the sparse inversion is the potential of energy leakage between the component shots that need to be deblended. In this paper we identify leakage as a basis misidentification problem and provide a reweighted thresholding method to reduce the leakage. Further, our study of the iterative thresholding and subtraction class of methods for source separation, indicate that existing model update procedures are suboptimal. We propose an updating step based on orthogonalization that has strong theoretical guarantees for improved convergence and is potentially more robust to leakage issues.

Introduction

A simultaneous source (SimSrc) acquisition can be represented using Berkhout's (2008) formulation, $\boldsymbol{d} = \boldsymbol{\Gamma m}$, where $\boldsymbol{\Gamma}$ is the blending matrix, $\boldsymbol{d}$ is the recorded data in $\boldsymbol{R}^N$ and $\boldsymbol{m}$ is our targeted unblended records in $\boldsymbol{R}^M$ with $M>N$. Since the system is underdetermined, methods have been proposed in the geophysics literature to solve this problem by adding constraints. The constraint used is the assumption that the model is sparse in a transform domain and the problem is rewritten as:

$$\boldsymbol{d} = \boldsymbol{\Gamma \Psi x} \quad \boldsymbol{m} = \boldsymbol{\Psi x} \tag{1}$$

where $\boldsymbol{\Psi}$ is the suitable transform domain where the model has a sparse representation, $\boldsymbol{x}$. Solving equation (1) by means of a sparse inversion is essentially similar to a rapidly emerging field known as compressive sensing (CS, Candes et al. 2006). From a CS standpoint existing SimSrc separation methods can be broadly classified as: (1) Convex relaxation: l_1 minimization (Moore et al. 2008, Akerberg et al. 2008, Lin and Herrmann 2009, Li et al 2013) and (2) Greedy algorithms: some form of iterative thresholding (Abma 2010, Doulgeris et al. 2011, Mahdad et al., 2011, Chen et al. 2013). The class (2) methods are usually easier to implement and faster than the class (1) methods, even though l_1 minimization when done in a suitable domain (e.g. curvelets) has stronger theoretical guarantees for convergence and model recovery. Our motivation for this paper is twofold: (1) Analyze the problem of leakage of energy between sources from a CS perspective and provide a simple way of trying to minimize the leakage and (2) Provide a method that improves the convergence behavior of greedy algorithms and can further minimize leakage.

Analysis of Leakage for SimSrc separation

For SimSrc separation leakage may be defined as the cross contamination of energy between the targeted deblended records. In this paper we analyze leakage as being caused by an incorrect support (and basis) identification of the model in its sparse domain, $\boldsymbol{\Psi}$. All sparse inversion methods have the underlying assumption that the model satisfies $\|\boldsymbol{x}\|_0 \leq \boldsymbol{s}$, and is thus *s-sparse* (at most $\boldsymbol{s}$ significant model components or *support* that can fully explain the model space) with $\boldsymbol{s} \ll \boldsymbol{M}$ in some suitable transform domain. The goal of the inversion is to correctly identify this support and the corresponding basis for the model space. Whenever this is incorrect we can expect leakage issues to show up.

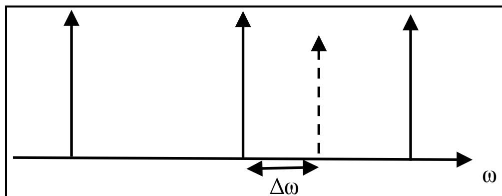


Figure 1: 1D representation of error in basis selection. Solid lines are the true basis vectors for the sparse signal. Dashed line is the 2nd basis identified incorrectly with an offset Δω to the true basis.

We develop the analysis based on greedy methods in the Fourier basis, though the theory applies to all general methods in any transform domain for sparse inversion of the SimSrc problem. Consider a 1D signal that is sparse in the frequency domain and has its basis vectors denoted by the bold arrows in Figure (1), located at their support locations. The thresholding step in the Fourier space can be defined as the set $\boldsymbol{J}$ containing the support values $\boldsymbol{a_j}$ (most coherent spectrum values) at the current iteration as:

$$\boldsymbol{J} = \{\boldsymbol{a_j}\} = |\langle \boldsymbol{v_i}, \boldsymbol{R_{k-}} \rangle| \geq \boldsymbol{t_s} \tag{2}$$

where $\boldsymbol{R_k}$ is the residual at the $\boldsymbol{k^{th}}$ iteration $\boldsymbol{t_s}$ is the threshold value, $|\langle \boldsymbol{q,m} \rangle|$ is the absolute value of the vector inner product or correlation and $\boldsymbol{v_i}$ is a candidate basis vector in the Fourier domain. The support is thus identified as the row vector from the Fourier kernel forming the smallest angle with the residual. When the support is misidentified the corresponding basis vector would show up at on offset to the true vector (dashed line in Figure 1) and/or as a damped version of true basis. Let $\boldsymbol{\Psi_t}$ be the true basis where the signal has a sparse representation and let $\boldsymbol{\Psi_b}$ be the estimated basis

DOI http://dx.doi.org/10.1190/segam2014-1226.1

which has at least one suboptimal basis vector selection. Following equation (1) our inversion scheme assumes the model to be sparse in the estimated incorrect basis, $\boldsymbol{\Psi_b}$ as:

$$\boldsymbol{m} = \boldsymbol{\Psi_b x_b} \quad (3)$$

While the model is actually sparse in the true basis as:

$$\boldsymbol{m} = \boldsymbol{\Psi_t x_t} \quad (4)$$

Consequently we have a resolution matrix that connects the true sparse estimate, $\boldsymbol{x_t}$ with the incorrectly estimated $\boldsymbol{x_b}$ via:

$$\boldsymbol{x_b} = \boldsymbol{\Psi^{-1}{}_b \Psi_t x_t} \quad (5)$$

We define the matrix $\boldsymbol{\Psi^{-1}{}_b \Psi_t}$ as the resolution matrix that defines the leakage in SimSrc separation. The resolution matrix would lead to smearing across components of the estimated signal producing leakage. Insertion of (5) into (1) gives:

$$\boldsymbol{d} = \boldsymbol{\Gamma \Psi_t \Psi^{-1}{}_b \Psi_t x_t} \quad (6)$$

Equation (6) explains the degradation in separation quality from a compressive sensing (CS) standpoint. One of the key assumption in CS for solving equation (1) is that the signal is sparse in the transform domain, i.e. its coefficients decay fast, via some power law. The resolution matrix weakens this assumption by smearing components across all model coefficients and thus the expected decay of the coefficients no longer happens. In other words $\boldsymbol{x}$ is no longer *s-sparse*. Notice that the smearing not only happens across the targeted unblended records, but also within each record producing undesirable noise in the final separation. The problem arises essentially because the support of a frequency sparse signal are themselves sparse only when its FFT/DFT Fourier coefficients are exact integer multiples of the Fourier basis' fundamental frequency (Duarte et al. 2010), which is generally not the case.

Notice the similarity with Anti-leakage Fourier transform (ALFT) based data regularization (Xu et al. 2005) where the resolution matrix in equation (6) now governs the well-known spectral leakage issue. ALFT-regularization can thus be classified as a sparse greedy inversion where the blending matrix is simply a sampling operator and the inversion imposes a sparsity constrain to reconstruct the data on the intended regular grid. We can also see that a potential problem with iterative thresholding based regularization like ALFT-regularization and POCS interpolation (Abma 2006), when the missing data has a regular pattern can be explained when viewed from a sparse CS inversion framework. The regular pattern in the missing traces imposes a regular pattern in the blending matrix, and thus one of the key requirements for the inversion to work effectively that the blending matrix and the sparse basis need to be maximally incoherent is weakened.

Compared to data regularization, the SimSrc problem has an added level of complexity, introduced by recording the data at a close to sub Nyquist rate. The practical implication of this is that we should expect separation to be much more sensitive to acquisition design parameters (particularly spacing and randomness in the shooting). We note that the analysis done in this section opens up the exciting possibility of combining deblending and regularization into one step. But this issue is not explored further in this paper.

Reweighted thresholding

A technique commonly used in thresholding processes for data regularization and noise suppression (Qin et al. 2012) is to compute radial weights for the spectrum in an unaliased band. The idea is to use these weights to enhance coherent energy and also suppress aliased events which do not start from the origin. Based on equation (2) we can say that such weighting improves the support identification. In the SimSrc problem, weights are computed in a spectrum that has contribution from multiple sources. Such weights tend to have a high degree of contamination compared to the data regularization problem. Thus a simple strategy to improve the support identification step is using a reweighting scheme for the spectrum. Once a first pass of

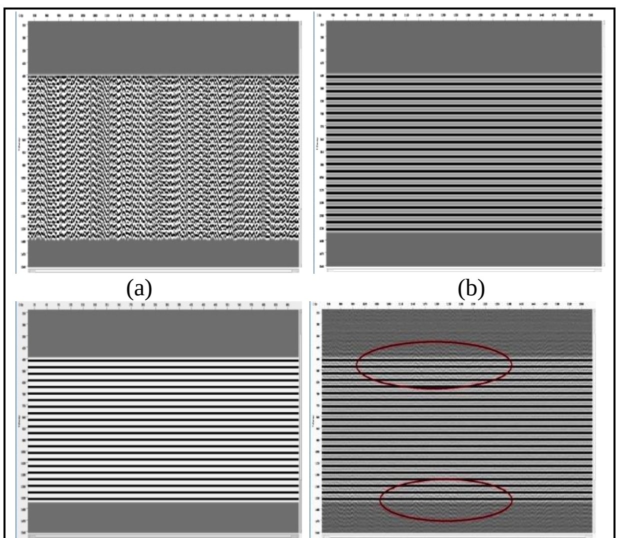
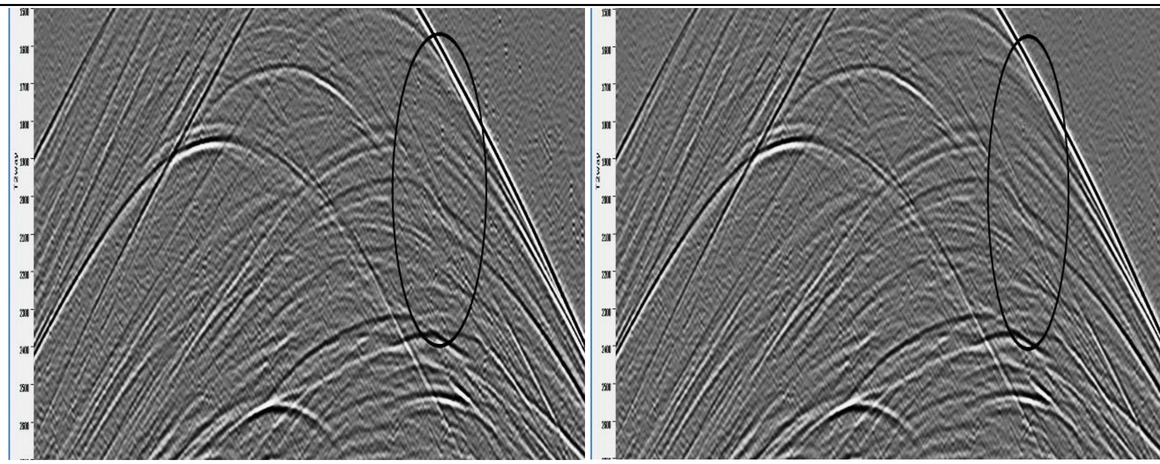


Figure 2: (a) blended data, (b) S1 estimate, (c) S2 estimate (5x stronger), (d) S1 estimate, without StOMP, at same iteration number. Notice the ringing noise left on the top and bottom in 2(d) indicated by the red ovals.

separation is complete, we recompute the weights on the initial separated outputs and then restart the iteration with the new weights. Since the recomputed weights are more robust,

DOI http://dx.doi.org/10.1190/segam2014-1226.1



(a) (b)

Figure 3:(a) Separation result without reweighting scheme and orthogonal projections on a receiver gather, (b) separation with reweighting and orthogonal projections, notice the leakage energy that is left within the black oval in (a) compared to (b), seen as "dots". Also note the over-all reduction in noise level

we expect thresholding or support identification for the second stage of iteration to be better. To effectively use the reweighting scheme a projection of the original blended data is done onto the model space so that each individual component of the model space (i.e. each record to be unblended) can have their own independent weights for the thresholding step.

Orthogonal Projections

The greedy methods when used for SimSrc separation can be generalized to be similar to a class of methods collectively known as Projection Pursuit (Huber, 1985) in the statistics community. An issue with convergence as well as potential leakage can occur (Donoho, 1985) if a particular projection is non-orthogonal. When viewed from such a projection pursuit framework, our implementation uses the following main projections: (1) Projection of blended data onto the model space, (2) Projection onto the support (thresholding), (3) Projection onto the feasible set (sparse model space), (4) Projection back onto the data space to update residual. Existing thresholding methods use projections (3) and (4) in the form:

$$\boldsymbol{m_k = m_{k-1} + A^T_{k-1} d} \qquad (7)$$

$$\boldsymbol{R_k = R_{k-1} - A_{k-1}\, m_k}\ , \qquad (8)$$

where $\boldsymbol{A = \Gamma \Psi_{k-1}}$, with the columns of $\boldsymbol{\Psi_{k-1}}$ populated with the vectors obtained during the thresholding step described in equation (2). Thus for each iteration it is guaranteed that the residual is orthogonal only to the basis vectors selected at the current iteration. However full backward orthogonality with *all basis vectors selected till* the current iteration is not maintained. We conclude that using equation (7) is a suboptimal estimate of the model. If a suboptimal model is used to update the residual, then support identifications for the next thresholding iteration could be suboptimal as well which might add to leakage. A natural solution is to enforce orthogonality between the residual and the column space of equation (1). For CS, such schemes have been proposed previously (Stagewise Orthogonal matching Pursuit, StOMP, Donoho et al., 2006) which we adopt for the SimSrc problem. Using orthogonal projections we update the model and the residual as:

$$\boldsymbol{m_k = (B^T_{k-1} B_{k-1})^{-1} B^T_{k-1} d} \qquad (9)$$

$$\boldsymbol{R_k = d - B_{k-1} m_k}\ , \qquad (10)$$

where $\boldsymbol{B = \Gamma \Psi_I}$, with the columns of $\boldsymbol{\Psi_I}$ populated with *all the basis vectors selected till* the k^{th} iteration. Another advantage of this is improved convergence (Tropp 2004). A similar orthogonal step was proposed for the data regularization problem (Hollander et al., 2012) but our implementation has two important distinctions: (1) StOMP allows multiple terms to enter the thresholding step, (2) We use the FFT instead of the DFT leading to faster implementation.

Examples

In Figure (2a) we show a simple synthetic where two sets of linear events (S1 and S2, S2 being 5 times stronger than S1) are combined together by applying small random time shifts to one of the datasets. We show the separation result of stopping the iterations after a fixed number of steps for S1, with (2b) and without orthogonal projections (2d). Notice the faster convergence for StOMP while providing perfect separation when iterations are stopped after a fixed number of steps. Due to the simplicity of the synthetic no reweighting schemes were used for this example.

DOI http://dx.doi.org/10.1190/segam2014-1226.1

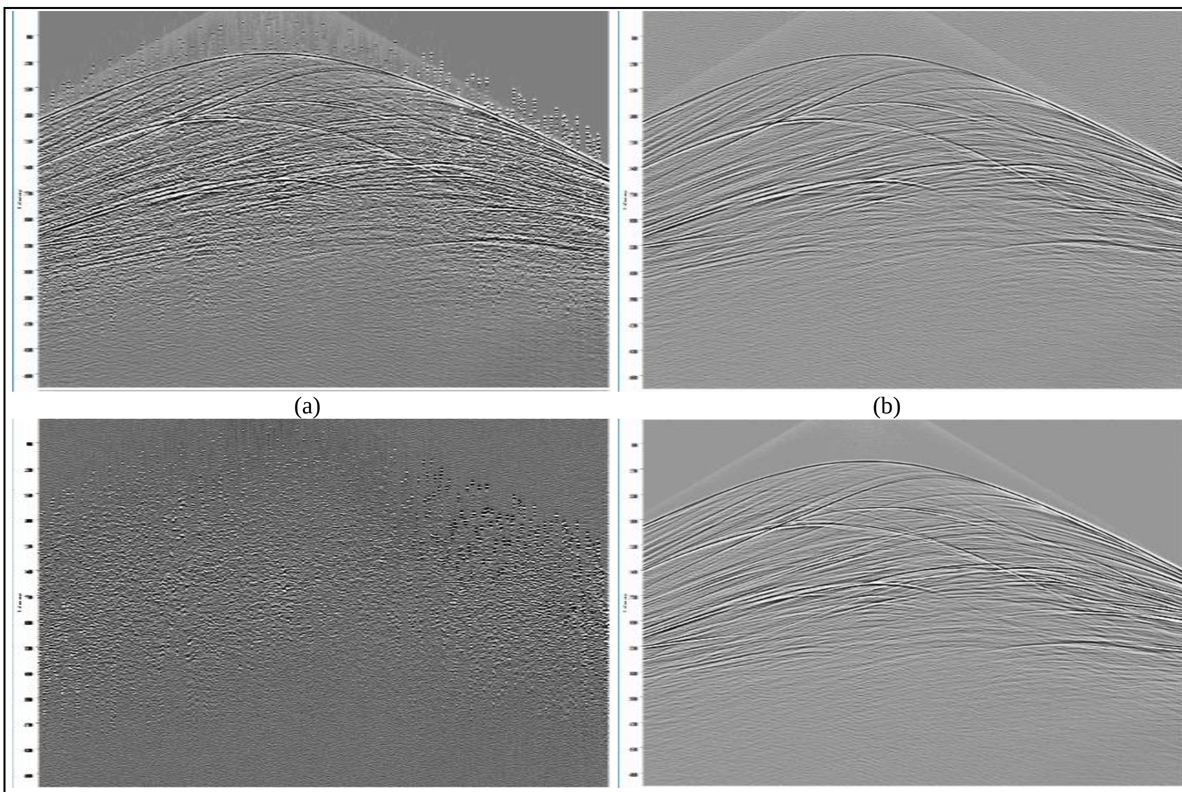
We now show results for a SimSrc data set simulated using the Marmousi synthetic. The data are generated by simply combining two records with a random time delay so that no continuous recording or random sampling (e.g. jittered sampling) of the records is done. Thus we test the sparse inversion in one of its worst case scenarios. Figure (3) shows a zoomed in comparison of the separation results without using reweighting and orthogonal projections (3a) and Figure (3b) shows the result of using both reweighting and orthogonal projections. Note the reduction of leakage energy within the black oval in (3b) compared to (3a). Finally Figure (4) shows the separation results on one receiver gather. The difference plot Figure (4c) shows little leakage of coherent energy and mostly contains the blending noise. This indicates that the separation (Figure 4b) is of high quality. Weak events are also well preserved when we compare the separation result (Figure 4b) with the original unblended data (Figure 4d).

Conclusions

In this paper we have defined the leakage problem in SimSrc separation as the result of incorrect support identification. Using this we have introduced a reweighting scheme for the spectrum to improve the support identification during the thresholding step. We have also used orthogonal projections for the model updating step to improve convergence and further reduce chances of potential leakage for the iterative thresholding methods. Our method produces good separation even in the case where the acquisition parameters do not allow the best utilization of the power of the CS-sparse inversion. We note that for successful and optimal source separation of field data substantial burden rests on the acquisition step as the problem is particularly sensitive to data quality.

Acknowledgements

We thank Manhong Guo and Will Whiteside for insightful discussions regarding high dimensional regularization algorithms. We thank Jian Mao for the Marmousi synthetic model. We also thank the TGS management for permission to publish this work.



(a) (b) (c) (d)

Figure 4: (a) blended data, (b) separated estimate, (c) difference between b and a, (d) original unblended data. Notice that very little coherent energy leaks into the difference plot in (c). Weak events, are well preserved when (b) and (d) are compared.

DOI http://dx.doi.org/10.1190/segam2014-1226.1

http://dx.doi.org/10.1190/segam2014-1226.1

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2014 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Abma, R. L., and N. Kabir, 2006, 3D interpolation of irregular data with a POCS algorithm: Geophysics, **71**, no. 6, E91–E97, http://dx.doi.org/10.1190/1.2356088.

Abma, R. L., T. Manning, M. Tanis, J. Yu, and M. Foster, 2010, High-quality separation of simultaneous sources by sparse inversion: 72nd Annual International Conference and Exhibition, EAGE, Extended Abstracts, B003.

Akerberg, P., G. Hampson, J. Rickett, H. Martin, and J. Cole, 2008, Simultaneous source separation by sparse Radon transform: 78th Annual International Meeting, SEG, Expanded Abstracts, 2801–2805.

Berkhout, A. J., D. J. Verschuur, and G. Blacquiere, 2008, Processing of blended data: Presented at the 78th Annual International Meeting, SEG.

Candès, E. J., J. K. Romberg, and T. Tao, 2006, Stable signal recovery from incomplete and inaccurate measurements: Communications on Pure and Applied Mathematics, **59**, no. 8, 1207–1223, http://dx.doi.org/10.1002/cpa.20124.

Chen, Y., S. Fomel, and J. Hu, 2013, Iterative deblending of simultaneous-source seismic data using shaping regularization: Presented at the 83rd Annual International Meeting, SEG.

Duarte, M. F., and R. G. Baranuik, 2010, Recovery of frequency sparse signals from compressive measurements: 48th Annual Conference on Communication, Control and Computing (Allerton), 599–606.

Donoho, D., I. Johnstone, and P. Rousseeuw, 1985, Discussion following article by P. Huber: Annals of Statistics, **13**, no. 2, 496–500, http://dx.doi.org/10.1214/aos/1176349526.

Donoho, D., Y. Tsaig, I. Drori, and J. L. Starck, 2012, Sparse solution of under determined systems of linear equations by Stagewise orthogonal matching pursuit: IEEE Transactions on Information Theory, **58**, no. 2, 1094–1121, http://dx.doi.org/10.1109/TIT.2011.2173241.

Doulgeris, P., A. Mahdad, and G. Blacquire, 2010, Separation of blended data by iterative estimation and subtraction of interference noise: 80th Annual International Meeting, SEG, Expanded Abstracts, 3514–3518.

Hollander, Y., D. Kosloff, Z. Koren, and A. Bartana, 2012, Seismic data interpolation by orthogonal matching pursuit: 74th Annual International Conference and Exhibition, EAGE, Extended Abstracts, B023.

Huber, P. J., 1985, Projection pursuit: Annals of statistics, **13**, no. 2, 435–475, http://dx.doi.org/10.1214/aos/1176349519.

Li, C., C. C. Mosher, L. C. Morley, Y. Ji, and D. Brewer, 2013: Joint source deblending and reconstruction for seismic data: 83rd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2013-0411.1.

Lin, T. Y., and F. J. Herrmann, 2009, Unified compressive sensing framework for simultaneous acquisition with primary estimation: 79th Annual International Meeting, SEG, Expanded Abstracts, 3113–3117.

Mahdad, A., P. Doulgeris, and G. Blacquiere, 2011, Separation of blended data by iterative estimation and subtraction of blending interference noise: Geophysics, **76**, no. 3, Q9–Q17, http://dx.doi.org/10.1190/1.3556597.

Moore, I., W. Dragoset, T. Ommundsen, D. Wilson, C. Ward, and D. Eke, 2008, Simultaneous source separation using dithered sources: 78th Annual International Meeting, SEG, Expanded Abstracts, 2806–2809.

Qin, F., R. M. Burnstad, and P. C. Leger, 2012, An effective f-k domain random noise suppression technique applied to a land data set: Presented at the 82nd Annual International Meeting, SEG.

Tropp, J. A., 2004, Greed is good: Algorithmic results for sparse approximation: IEEE Transactions on Information Theory, **50**, no. 10, 2231–2242, http://dx.doi.org/10.1109/TIT.2004.834793.

Xu, S., Y. Zhang, D. Pham, and G. Lambare, 2005, Antileakage Fourier transform for seismic data regularization: Geophysics, **70**, no. 4, V87–V95.

DOI http://dx.doi.org/10.1190/segam2014-1226.1