Simultaneous source separation using a robust Radon transform

Amr Ibrahim[*] and Mauricio D. Sacchi, University of Alberta

SUMMARY

The robust Radon transform is presented as an alternative to eliminate incoherent noise that surges in simultaneous source processing. The robust Radon transform is designed by utilizing an L_1 misfit that permits to estimate data predictions that are invulnerable to erratic noise. Seismic data obtained by simulations source techniques contain erratic noise in common-receiver domain. This presentation examines the removal of erratic noise in common-receiver domain via the robust Radon transform.

INTRODUCTION

Simultaneous acquisition have been proposed to reduce the acquisition cost of seismic surveys. However, most seismic processing techniques are designed to handle data with non-overlapping sources. Therefore, the separation of blended sources into their equivalent non-overlapped sources is an important step prior to classical processing sequences. Blending is equivalent to time-shifting each individual source data and summing according to certain scheme (encoding). Blending can be expressed mathematically as

$$\mathbf{b} = \Gamma \mathbf{d} \tag{1}$$

where $\mathbf{b}$ is the blended data, $\mathbf{d}$ indicates the data cube we would have liked to record in the absence of source overlapping and Γ is the blending operator. The blending operator in the frequency domain can be expressed via the following expression

$$[\Gamma]_{ij} = e^{i\omega\tau_{ij}} \tag{2}$$

where τ_{ij} is the delay of i-th source firing time with respect to the detector j. In order to compute the de-blended data one can follow two different approaches. In the first category of methods, deblending is posed as an inverse problem where one minimizes a cost function that includes a data misfit and a regularization term. In addition, rather than inverting directly for $\mathbf{d}$ one must invert for the representation of $\mathbf{d}$ in terms of coefficients $\mathbf{c}$ in an auxiliary domain. In other words, if the data are represented in terms of coefficients $\mathbf{c}$ in a basis Φ, such that $\mathbf{d} = \Phi\mathbf{c}$, the goal is to estimate $\mathbf{c}$ by minimizing the following cost function

$$J = \|\mathbf{b} - \Gamma\Phi\mathbf{c}\|_2^2 + \mu\mathscr{R}(\mathbf{c}) \tag{3}$$

The regularization term $\mathscr{R}(\mathbf{c})$ is needed because Γ is a non-invertible operator. Algorithms in this category include sparse Radon inversion (Moore et al., 2008; Akerberg et al., 2008) , iterative $f-k$ filtering (Mahdad et al., 2011; Doulgeris et al., 2012) and curvelet-based source separation (Wason et al., 2011; Lin and Herrmann, 2009). The inversion process can also be posed via a projected gradient optimization algorithm (Abma et al., 2010). All these methods attempt to retain coherent signal in common receiver gathers by imposing simplicity (sparsity) in $\mathbf{c}$.

An alternative strategy is to estimate $\mathbf{d}$ directly from the pseudo-deblended

$$\widetilde{\mathbf{d}} = \Gamma^H \mathbf{b} \tag{4}$$

via denoising methods (Huo et al., 2012; Berkhout, 2008). Pseudo de-blending is equivalent to applying shifts and dividing long blended records in those that one would have obtained via standard non-overlapping acquisitions. However, pseudo de-blending does not remove interferences resulting from the overlapping of different sources and deblended records contain a considerable amount of interferences. Interferences are coherent in common source gathers and incoherent in common receiver gathers.

In this work we propose to use the robust Radon transform (Ji, 2006, 2012) to eliminate incoherent noise in common receivers gathers that were obtained by application of the pseudo deblending operator (Equation 4) to simultaneous acquisition data.

RADON TRANSFORMS

In our analysis we consider data organized in common receiver gathers that have been obtained via pseudo-deblending. To avoid notational clutter, we will designate these data in common receiver gathers as $d(t,h)$ or in vector form $\mathbf{d}$ with t time and h source-receiver distance. Radon transforms in both frequency and time domains have been utilized to model seismic reflections and to attenuate coherent noise. For instance, the parabolic Radon transform (Hampson, 1986) has been widely used for multiple suppression in common mid point gathers after normal moveout correction correction. Similarly, Hyperbolic Radon transforms have been used to attenuate multiples in uncorrected common mid point gathers. Radon transforms with apex terms have been proposed to attenuate diffractions (Trad et al., 2003) and to reconstruct seismic gathers (Hokstad and Sollie, 2006) for 3D surface related multiple elimination. Many studies have proposed to improve the resolution and computational efficiency of both frequency domain and time domain Radon transforms. Thorson and Claerbout (1985) were the first authors to cast Radon transform as an inversion problem. They also proposed a sparse inversion method to obtain highly focused Radon gathers in the time domain. The original frequency domain parabolic Radon transform proposed by Hampson (1986) was modified by Sacchi and Ulrych (1995) to incorporate sparsity as well. robust and high resolution Radon transforms have been explored by Ji (2006, 2012).

Transforms that map images and signals into new domains like the Radon transform is a subject of intense research. In general, data are transformed to a new domain to facilitate separation of its components and to differentiate signals from noise. The Radon transform is an integral transform that in its dis-

DOI http://dx.doi.org/10.1190/segam2013-0447.1

crete form can be expressed via the following two expressions

$$\widetilde{m}(\tau,\xi_j) = \sum_h d(t = \widetilde{\phi}(\tau,h,\xi),h)\,, \tag{5}$$

$$d(t,h) = \sum_\xi m(\tau = \phi(t,h,\xi),\xi)\,, \tag{6}$$

where $d(h,t)$ denotes the data (common receiver gather in this article) , $\widetilde{m}(q,\xi)$ is Radon coefficients that one can obtain using the adjoint Radon operator. The parameter ξ is the Radon parameter that depends on the type of integration path that one adopts for the Radon integral. Popular alternatives are the linear, parabolic and apex-shifted hyperbolic transforms with travel-time integration paths provided in table (1). Equations (5) and (6) define linear operators that can be represented using the language of linear algebra. For instance, the action of the adjoint Radon operator on the data can be expressed via matrix times vector multiplication

$$\widetilde{\mathbf{m}} \quad = \mathbf{L}^T \mathbf{d} \tag{7}$$

similarly, the action of the forward Radon operator on the model $\mathbf{m}$ can represented by

$$\mathbf{d} \quad = \mathbf{L}\mathbf{m}\,. \tag{8}$$

In general, we use equation (8) to estimate $\mathbf{m}$ via an inversion procedure and then, the estimated Radon model is used to reconstruct an improved version of $\mathbf{d}$.

Table 1: Radon operators: HRT: Hyperbolic Radon ransform. ASHRT: Apex Shifted Hyperbolic Radon Transform. PRT: Parabolic Radon Transform. ASPRT: Apex Shifted Parabolic Radon Transform. LRT: Linear Radon Transform. h: Offset. v: Velocity. h_0: Apex. q: Curvature . p: dip.

Operator	ξ	$\phi(\tau,h,\xi)$	$\widetilde{\phi}(t,h,\xi)$
HRT	$\xi = v$	$t = (\tau^2 + \frac{h^2}{v^2})^{1/2}$	$\tau = (t^2 - \frac{h^2}{v^2})^{1/2}$
ASHRT	$\xi = [v,h_0]$	$t = (\tau^2 + \frac{(h-h_0)^2}{v^2})^{1/2}$	$\tau = (t^2 - \frac{(h-h_0)^2}{v^2})^{1/2}$
PRT	$\xi = q$	$t = \tau + qh^2$	$\tau = t - qh^2$
ASPRT	$\xi = [q,h_0]$	$t = \tau + q(h-h_0)^2$	$\tau = t - q(h-h_0)^2$
LRT	$\xi = p$	$t = \tau + ph$	$\tau = t - ph$

ROBUST INVERSION

We assume data contaminated with noise and therefore, we pose the estimation of $\mathbf{m}$ via the reduction of the residuals

$$\mathbf{r} = \mathbf{d} - \mathbf{L}\mathbf{m}\,. \tag{9}$$

The problem can be formulated in terms of the minimization of a cost function given by

$$J = \|\mathbf{r}\|_p^p + \mu \|\mathbf{m}\|_q^q. \tag{10}$$

where the first term on the right hand side indicates the misfit term whereas the second term is used to indicate the regularization term. By minimizing the cost function with respect to the unknown vector of Radon coefficients $\mathbf{m}$ one finds a solution that reproduces the observations $\mathbf{d}$. The regularization term is used to impose desired characteristics to the unknown vector of Radon coefficients. The parameters p and q ($1 \le q,p \le 2$) represent exponent of the p-norm of the misfit and the $q-$norm of the model regularization term, respectively. In essence, we adopt $p = 2$ for data contaminated with Gaussian noise. On the other hand, we prefer $p = 1$ when the data are contaminated by erratic noise. The regularization term can be used to retrieve minimum norm solutions when $q = 2$ or sparse solutions when $q = 1$. We also represent the p (and q) norm by the following expression

$$\|\mathbf{x}\|_p^p = \sum_i |x_i||x_i|^{p-2}|x_i| = \|\mathbf{W_x}\,\mathbf{x}\|_2^2 \tag{11}$$

where $[\mathbf{W_x}]_{ii} = \frac{1}{\sqrt{|x_i|^{2-p}}}$.

The norm is analytically correct and valid for $x_i = 0$. However, our algorithm will treat the variable x_i in the denominator and that on the numerator as two independent variables during the iterative process adopted to minimize J. Therefore, we modify the matrix of weights for $\mathbf{m}$ as follows

$$[\mathbf{W_m}]_{ii} = \begin{cases} \frac{1}{\sqrt{|m_i|^{2-p}}} & \text{if } m_i > \varepsilon_m \\ \frac{1}{\sqrt{\varepsilon^{2-p}}} & \text{if } m_i \le \varepsilon_m\,. \end{cases} \tag{12}$$

Similarly, we define the matrix of weights for the residual as follows

$$[\mathbf{W_r}]_{ii} = \begin{cases} \frac{1}{\sqrt{|r_i|^{2-q}}} & \text{if } r_i > \varepsilon_r \\ \frac{1}{\sqrt{\varepsilon^{2-q}}} & \text{if } r_i \le \varepsilon_r\,. \end{cases} \tag{13}$$

With the latter in mind, we can now replace the non-quadratic optimization problem (Equation 10) by a sequence of quadratic optimization problems where we minimize the cost function

$$J' = \|\mathbf{W}_r^p\,\mathbf{r}\|_2^2 + \mu \|\mathbf{W}_m^q \mathbf{m}\|_2^2. \tag{14}$$

The constants ε_m and ε_r represent the transition from an L_2 to an L_1 minimization problem. Holland and Welsch (1977) used robust statistics to estimate the value of ε_r

$$\varepsilon_r = b_r \ \frac{MAD(\mathbf{r})}{0.6745}, \tag{15}$$

where MAD indicates the median absolute deviation of the residuals $\mathbf{r}$. The parameter b_r is a tuning parameter. Holland and Welsch (1977) recommended using $b_r = 1.345$. The parameter ε_m is computed via the following expression

$$\varepsilon_m = b_m \ \frac{max(\mathbf{m})}{100}$$

where b_m is a tuning parameter that in our simulations was selected in an heuristic fashion and set to $b_m = 0.5$.

Expression (14) is minimized via the method of conjugate gradients followed by an update of the matrices of weights $\mathbf{W_r}$ and $\mathbf{W_m}$ (Trad et al., 2003). In essence, we have an internal iteration to minimize Equation (14) via the method of conjugate gradients and an external iteration to update the weights. The method of conjugate gradient is stopped when the change in misfit between iterations is less than a pre-defined tolerance value ($tolerance = 0.01$). The external group of iterations was set to 5. Notice that when $p = 2$ and $q = 2$ we have the classical least-squares Radon transform with quadratic regularization that can be directly solved with the method of conjugate

DOI http://dx.doi.org/10.1190/segam2013-0447.1

gradients because $\mathbf{W_m}$ and $\mathbf{W_r}$ are identity matrices. When $p = 1$ and $q = 2$ we have the high resolution (sparse) Radon transform (Trad et al., 2003).

EXAMPLES

We tested the robust Radon transform with a marine data set from the Gulf of Mexico. The data were numerically blended with 50% time reduction compared to the conventional acquisition. The source firing times versus source location are displayed in Figure 1. Four moving sources were used to simulate a practical blended acquisition survey. Each source fires while it is moving along the same line and in the same direction as other sources.

The original data utilized for the experiment is portrayed in Figure 2a. These data were numerically blended and pseudo-deblended to obtain Figure 2b. The data recovered by forward modelling with the Radon operator (apex shifted hyperbolic Radon operator) are shown in Figure 2c. In this example, the Radon transform model was individually estimated for each common receiver gather using the robust Radon transform ($p = 1$, $q = 2$). The difference cube in Figure 2d displays the reconstructed error.

The Radon model was estimated using least squares ($p = q = 2$, sparse ($p = 2$, $q = 1$) and robust inversion ($p = 1$, $q = 2$) and the resulting comparison of Radon panels (in time, apex, velocity space) are displayed in Figure 3. Clearly, in this test we used the apex shifted hyperbolic Radon transform because in common receiver domain one does not have a priori knowledge of the position of the apexes. Figure 4 shows the common receiver gather retrieved by each type of inversion after forward modelling the Radon panels in Figure 3. To quantitatively determine the accuracy of the data estimated by our processes we define,

$$Q = 10Log\frac{\|\mathbf{d}_{original}\|_2^2}{\|\mathbf{d}_{original} - \mathbf{d}_{retrieved}\|_2^2}. \tag{16}$$

The Q value for the recovered common receiver gather using least squares, sparse and robust transforms are are 10.71, 7.18 and 14.28, respectively. In addition, Figure 5 shows a close look at the results of Figure 4. These figures confirm that modelling the data via the robust Radon transform is an effective mean to remove interferences. The exercise also shows the importance of equipping the design of the Radon operator with a robust misfit functional. We were surprised to find out in our numerous tests that that robustness is more important than sparsity in this particular type of experiments. Solving the problem with $p = q = 1$ does not lead to better solutions than those obtained with $p = 1$ and $q = 2$. In fact, adding sparsity to the Radon solution in addition to robustness is not as simple as one might think because the algorithm becomes quite sensitive to the selection of ε_r and ε_m. Recent results in the area of robust deconvolution that include sparsity constraints (Gholami and Sacchi, 2012) suggest that more sophisticated algorithms are needed to obtain solutions that are not prone to failure due to incorrect parameter selection.

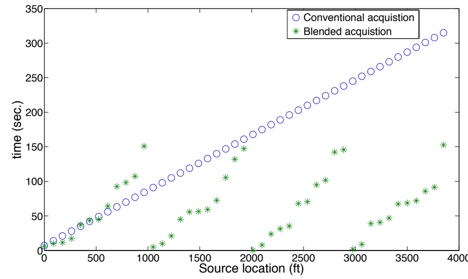

Figure 1: Sources firing times for a blending experiment.

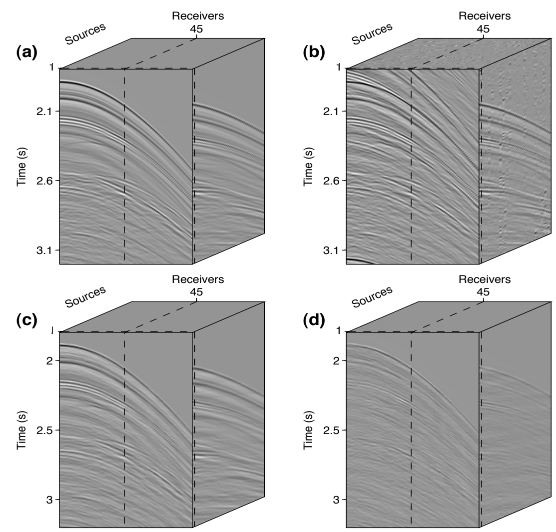

Figure 2: Real data example. (a) Original. (b) Pseudo deblended. (c) Data recovered by forward modelling with the robust Radon transform $p = 1$ and $q = 2$. (c) Difference between the estimated and the original common receiver gather.

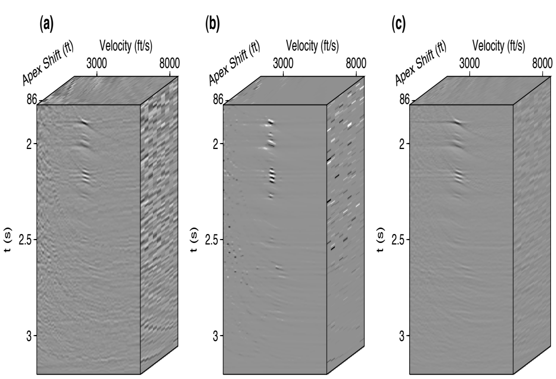

Figure 3: Radon panels obtained via inversion with (a) $p = 2$ and $q = 2$, (b) $p = 2$ and $q = 1$ and, (c) $p = 1$ and $q = 2$.

DOI http://dx.doi.org/10.1190/segam2013-0447.1

CONCLUSION

Interferences in common receiver gathers can be removed using robust Radon transforms. Treating blending noise as outliers and modelling the data with a Radon transform equipped with a robust misfit functional is an effective way of removing interferences caused by simultaneous shooting acquisition. Our tests showed that using a robust misfit ($p = 1$) with simple quadratic regularization ($q = 2$) produces better results than imposing sparsity on the model ($q = 1$).

ACKNOWLEDGEMENTS

The authors are grateful to the sponsors of Signal Analysis and Imaging Group (SAIG) at the University of Alberta.

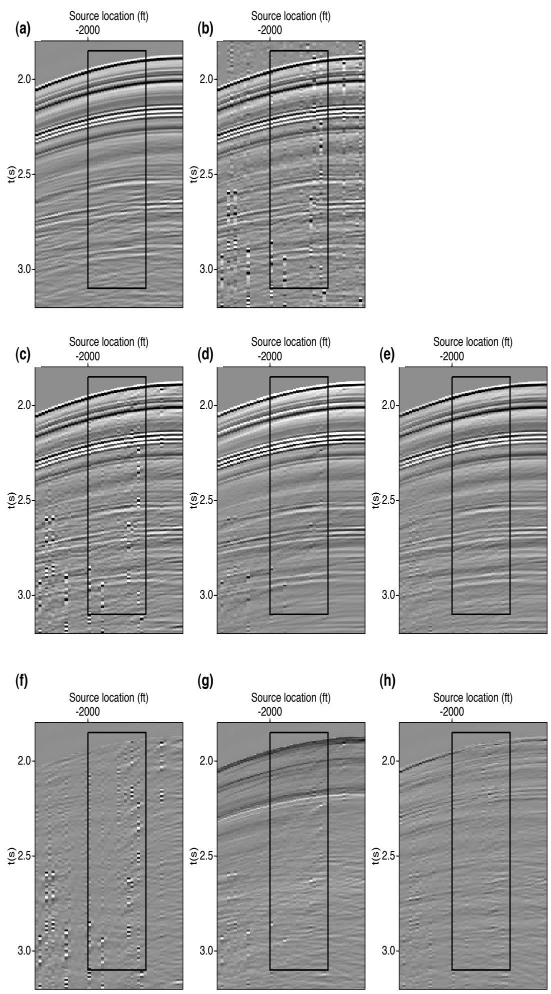

Figure 4: Common receiver gather. (a) Original data. (b) Blended data. (c) Data retrieved using $p = 2$ and $q = 2$. (d) Data retrieved using $p = 2$ and $q = 1$. (e) Data retrieved using $p = 1$ and $q = 2$. (f), (g) and h) Error displays for (c), (d) and (e), respectively.

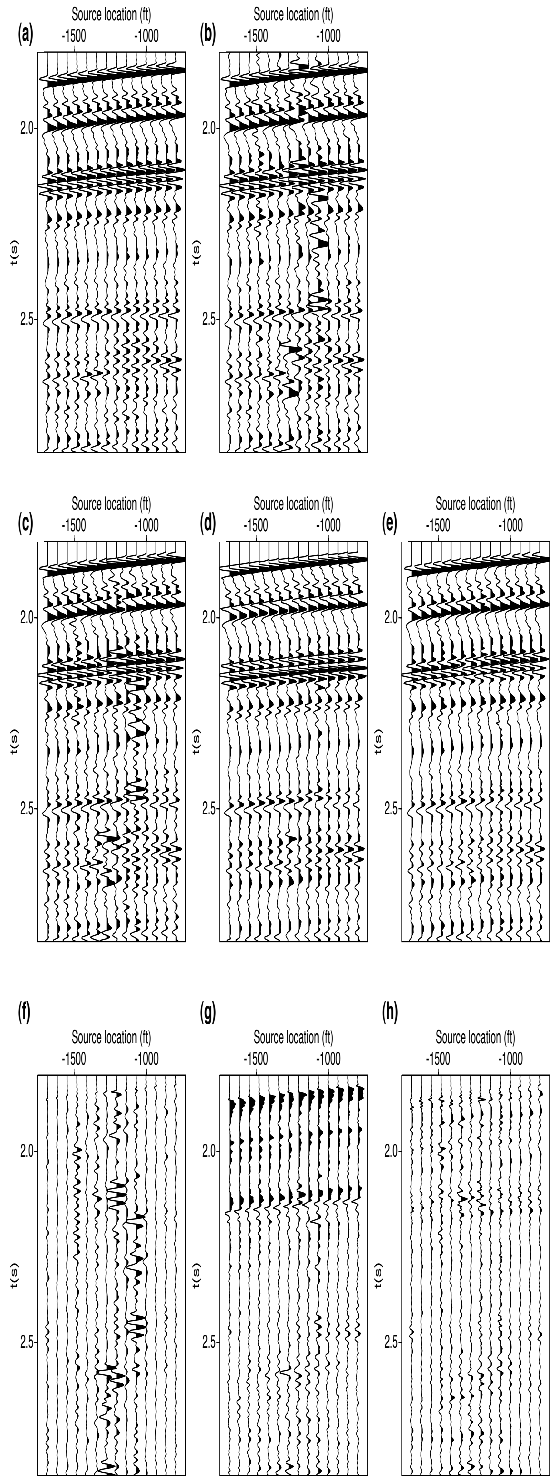

Figure 5: Close up of the real data example. (a) Original unblended data. (b) Blended data. (c) Data estimated using $p = 2$ and $q = 2$. (d) Data estimated using $p = 2$ and $q = 1$ (e) Data estimated using $p = 1$ and $q = 1$. (f), (g) and h) Error displays for (c), (d) and (e), respectively.

DOI http://dx.doi.org/10.1190/segam2013-0447.1

http://dx.doi.org/10.1190/segam2013-0447.1

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2013 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., T. Manning, M. Tanis, J. Yu, and M. Foster, 2010, High quality separation of simultaneous sources by sparse inversion: 72nd Conference and Exhibition, EAGE, Extended Abstracts, 109–111.

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., 2008, Changing the mindset in seismic data acquisition: The Leading Edge, **27**, 924–938, http://dx.doi.org/10.1190/1.2954035.

Doulgeris, P., K. Bube, G. Hampson, and G. Blacquière, 2012, Convergence analysis of a coherency-constrained inversion for the separation of blended data: Geophysical Prospecting, **60**, 769–781, http://dx.doi.org/10.1111/j.1365-2478.2012.01088.x.

Gholami, A., and M. D. Sacchi, 2012, A fast and automatic sparse deconvolution in the presence of outliers: IEEE Transactions on Geoscience and Remote Sensing, **50**, 4105–4116, http://dx.doi.org/10.1109/TGRS.2012.2189777.

Hampson, D., 1986, Inverse velocity stacking for multiple elimination: 66th Annual International Meeting, SEG, Expanded Abstracts, 422–424.

Hokstad, K., and R. Sollie, 2006, 3D surface-related multiple elimination using parabolic sparse inversion: Geophysics, **71**, no. 6, V145–V152, http://dx.doi.org/10.1190/1.2345050.

Holland, P., and R. Welsch, 1977, Robust regression using iteratively reweighted least-squares: Communications in Statistics Theory and Methods, **6**, 813–827, http://dx.doi.org/10.1080/03610927708827533.

Ji, J., 2006, CGG method for robust inversion and its application to velocity-stack inversion: Geophysics, **71**, no. 4, R59–R67, http://dx.doi.org/10.1190/1.2209547.

Ji, J., 2006, Robust inversion using biweight norm and its application to seismic inversion: Exploration Geophysics, **43**, 70–76.

Lin, T. T., and F. J. Herrmann, 2009, Designing simultaneous acquisitions with compressive sensing: 71st Conference and Exhibition, EAGE, Extended Abstracts, 23951.

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.

Moore, I., B. Dragoset, T. Ommundsen, D. Wilson, C. Ward, and D. Eke, 2008, Simultaneous source separation using dithered sources: 78th Annual International Meeting, SEG, Expanded Abstracts, **27**, 2806–2810, http://dx.doi.org/10.1190/1.3063928.

Sacchi, M., and T. Ulrych, 1995, High-resolution velocity gathers and offset space reconstruction: Geophysics, **60**, 1169–1177, http://dx.doi.org/10.1190/1.1443845.

Thorson, J. R., and J. F. Claerbout, 1985, Velocity-stack and slant-stack stochastic inversion: Geophysics, **50**, 2727–2741, http://dx.doi.org/10.1190/1.1441893.

Trad, D., T. Ulrych, and M. Sacchi, 2003, Latest views of the sparse Radon transform: Geophysics, **68**, 386–399, http://dx.doi.org/10.1190/1.1543224.

Wason, H., F. J. Herrmann, and T. T. Y. Lin, 2011, Sparsity promoting recovery from simultaneous data: A compressive sensing approach: 81st Annual International Meeting, SEG, Expanded Abstracts, 6–10, http://dx.doi.org/10.1190/1.3628174.

DOI http://dx.doi.org/10.1190/segam2013-0447.1