Sparse inversion based deblending in CMP domain using Radon operators

Kai Zhuang, Daniel O. Trad, and Amr Ibrahim

ABSTRACT

We implemented a deblending framework in the CMP domain to test the efficacy of deblending outside the commonly used receiver domain. By operating in the CMP domain instead of the common receiver domain, dipping reflectors are centered as opposed to apex shifted, this allows us to implement a simple hyperbolic Radon operator to decrease processing time taken to invert for a deblended data set versus an apex shifted operator. The Radon operator is posed as an inversion problem using a $\mathcal{L}_1$ model norm to support focusing in the Radon domain allowing better mapping of data back to their focused gathers. Implementation of deblending in an inversion-based framework is a relatively newer route to exploring deblending, with previous source separation implementations being denoising the pseudo-deblended data. Inversion based deblending allows us to explain all the data by refitting back to the blended dataset using the blending operator.

INTRODUCTION

Seismic exploration research is an economics driven field, and as a result, methods of cost reduction while retaining quality and resolution are often a point of interest. Simultaneously sourced acquisition originally proposed in 1983 (Garottu, 1983) and expanded upon in 1998 (Beasley et al., 1998), is a method to reduce the time spent in the field collecting data by firing multiple sources at the same time. This is in contrast to traditional acquisition which tries to avoid source interference. In blended acquisition, sources are allowed to interfere while they are later separated in processing thus reducing time spent in the field. There are generally two main streams of thought to deblending data, the denoising approach to remove blending noise (Mahdad et al., 2011), and the inversion-based approach (Abma et al., 2015) to refit the entire dataset. Both methods work best when the simultaneously fired shots are fired with a small randomized delay time between each other (Berkhout, 2008). Deblending takes advantage of random delay times of each source through a process called pseudo deblending. Pseudo deblending moves shots into their own data panels mimicking a conventional survey. After pseudo deblending is applied, the shots to be separated show up as coherent events while the other interfering shots show up as randomized blending noise. Deblending through the use of the blending operator (Berkhout, 2008) has originally be applied through the denoising method. Many different deblending approaches include direct separation in the shot domain using apex shifted Radon (Trad et al., 2012), the use of the Stolt operator as a denoising method in the receiver domain (Ibrahim and Sacchi, 2015) (Ibrahim and Sacchi, 2014), a thresholding approach in the FK domain using an inversion scheme (Stanton and Wilkinson, 2018), and a proposed deblending using migration-demigration operators (Trad, 2015). This paper introduces inversion based deblending in the CMP domain as a high-speed approach using a relatively simple Radon focusing operator.

THEORY

Blending operator

The blending operator allows us to transition between the blended domain where sources are overlapped, commonly known as a super shot and the unblended domain where sources belong to individual shot records. This can be represented as the blending operator $\mathbf{\Gamma}$ which contains the overlap and delay time information. We can then represent our unblended data as $\mathbf{D}$ and out blended data as $\mathbf{D_{bl}}$ with the following relationship (Berkhout et al., 2009; Berkhout, 2008)

$$\mathbf{D_{bl}} = \mathbf{\Gamma D}. \tag{1}$$

By implementing the adjoint of the blending operator on the blended data we can then recover the pseudo-deblended data, which is represented in the unblended space, with each shot having its own data frame as opposed to multiple shots per data frame. This can be expressed by the following (Berkhout et al., 2009; Berkhout, 2008)

$$\tilde{\mathbf{D}} = \mathbf{\Gamma^H D_{bl}}. \tag{2}$$

$\tilde{\mathbf{D}}$ represents the pseudo-deblended data in standard shot frames. The blending matrix serves only to add shots together with random delay times and as such it is just a superposition of the individual shot record's data. This also means that pseudo-deblending acts to redistribute the data to their individual frames. This forms an overcomplete representation of the data with redundant information stored in each and every pseudo-deblended source panel. Note that the adjoint operator does not remove the source interference from the data and only serves to copy it.

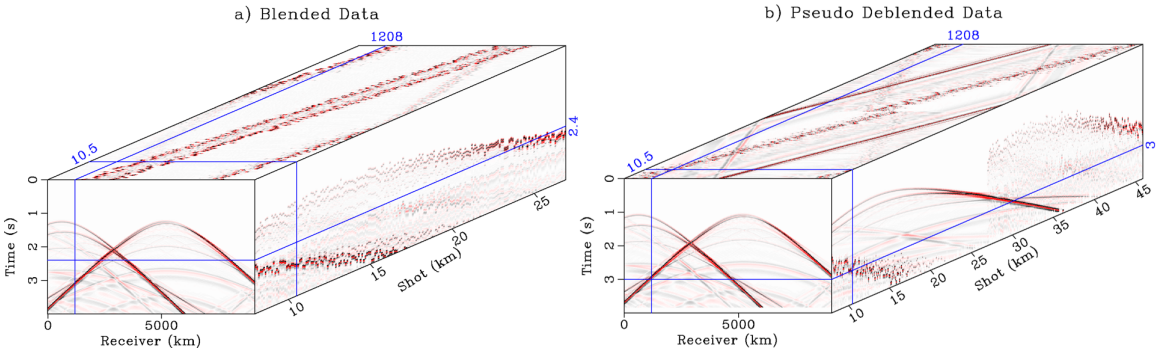

FIG. 1. Blended vs Pseudo deblended data: Blended data in a) and pseudo deblended data in b), the most significant change can be noted in the common reciever domain where a single shot becomes coherant per frame, the same effect occurs in offset and CMP domains.

One of the key factors to make deblending easier is the random delay incorporated into the blending scheduling, and in extension the blending operator. This random delay between shots becomes important when we apply the adjoint blending operator commonly referred to as pseudo-deblending to the blended data set. We can view the blending operation as an operation on the standard acquisition data, for our example we will call $\mathbf{S}$ our

acquisition data, seen in Figure 2. The columns of $\mathbf{S}$ represent each individual survey and the rows are each individual shot. We can then introduce the random delay times as just a linear phase shift in the frequency domain in the blending operator $\mathbf{\Gamma}$. This then gives us the blended data $\mathbf{S_{bl}}$ such that each survey, contains multiple shots each of which has a random time delay.

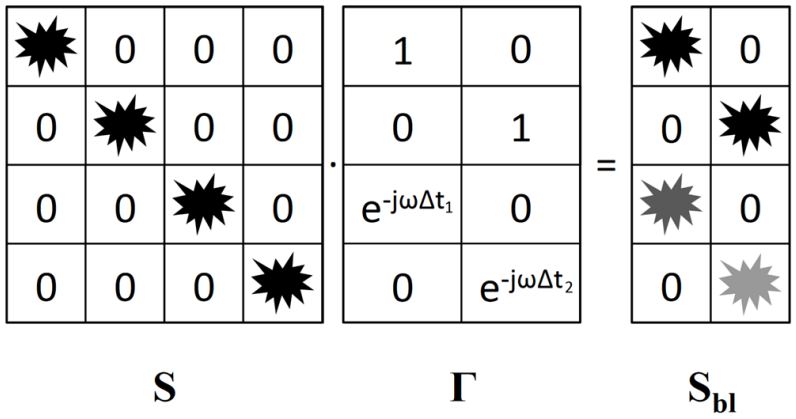

FIG. 2. Illustration of the Blending operator Γ (Urruticoechea, 2015): Where the unblended source is blended using the blending operator to output the blended data set.

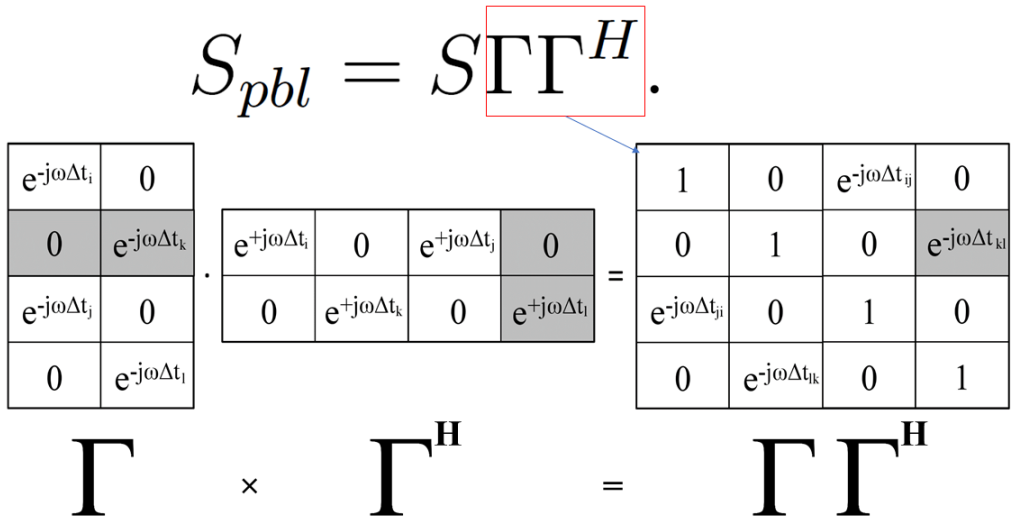

FIG. 3. Illustration of the effect of pseudo deblending (Urruticoechea, 2015): Where Γ is the blending operator and Γ^H is the pseudo deblending operator and combined is an operation on the true data S.

The adjoint of the blending operator is then represented as the conjugate transpose of the blending operator Γ^H which then can be combined with Γ to show the behavior of pseudo deblending seen in Figure 3.

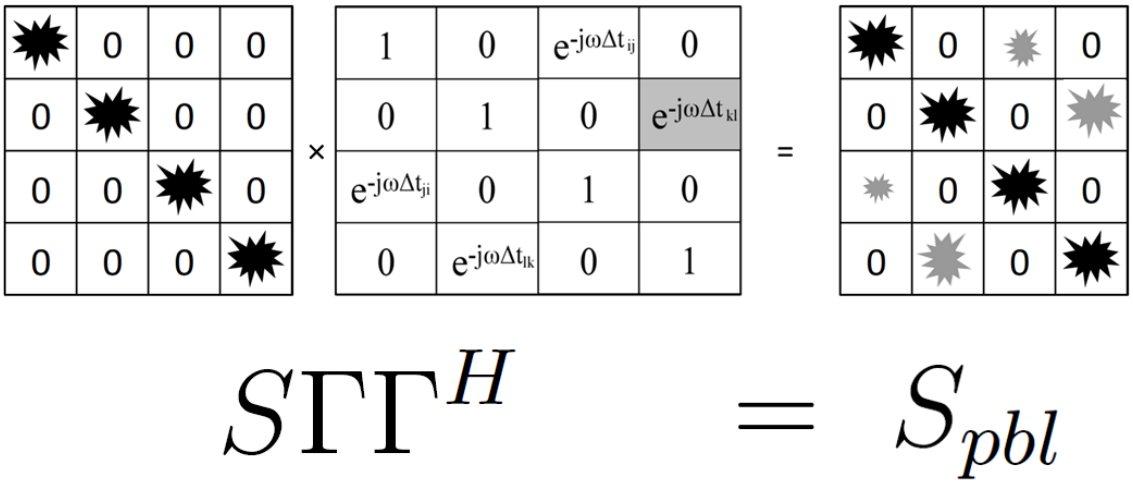

FIG. 4. Illustration of the effect of blending and pseudo deblending (Urruticoechea, 2015): It can be seen that the main data is preserved in the main diagonal and off diagonal events are phase shifted/time delayed.

With the random delay times, after pseudo-deblending the data, transformation into another domain like the CMP or receiver domain shows us that one set of events appear as coherent or in focus while other sets of events are incoherent or out of focus. The coherence of the pseudo deblended data pertains to shots that have their delay times corrected for in pseudo deblending. The coherent events correspond with the corrected shots, while the out of focus information corresponds to the shots that are not properly corrected, and thus does not belong in that particular shot frame. By utilizing a coherency based denoising or inversion algorithm, removal of the incoherent blending noise can be achieved leaving only the unblended coherent signal.

Radon Transform

Due to the forward and adjoint blending operators being very simple there is no unique solution to the inverse thus it is considered ill-posed, and as such a constraint is required to aid in the inversion of blended data. The hyperbolic Radon transform acts as a focusing operator representing a coherency constraint for the inversion of blended data. The hyperbolic Radon transform is represented by the equation (Thorson and Claerbout, 1985)

$$u(p,\tau) = \int_{h_1}^{h_2} d(h, t = \sqrt{\tau^2 + p^2 h^2}) dh. \tag{3}$$

Where $u(p,\tau)$ is the radon space data, p is the slowness, τ is the two way travel time, h1 is the upper offset limit, h2 the lower offset limit, and d is the data space to be transformed. The slowness p is then defined as the inverse of velocity $1/v$. This operator focuses events into points along hyperbolic arrivals according to a range of moveout velocities. The Radon transform has been used in many applications such as interpolation in frequency or time domain (Trad et al., 2003), velocity analysis (Thorson and Claerbout, 1985), and deblending via denoising (Ibrahim and Sacchi, 2015). The sparse hyperbolic Radon transform re-formats the ill-posed inversion to one that uses sparse regularization, with only a few

points in the radon space representing a large amount the information in shot space. By implementing a sparse constraint on the combined operators we can then implement an inversion based scheme to deblend the blended data. A limitation to this basic hyperbolic radon operator is that the operator only focuses events on a single apex location, generally a specified midpoint where the offset h is 0. The issue with operating at single apex locations is that in seismic data, dipping reflectors have their apexes shifted relative to zero offset and thus will not properly map into the radon domain, this comes as a significant issue due to seismic data commonly being complex a dipping. There are two solutions to this issue, either the implementation of an apex shifted hyperbolic radon transform (Trad et al., 2004), or to apply the hyperbolic Radon transform in a domain where apexes are not shifted. In this paper, the basic hyperbolic radon transform is implemented in the CMP domain where dipping events are centered (Claerbout, 1985) and thus the hyperbolic Radon transform will be able to map all events properly.

Sparse inversion

Sparse inversion seeks to find the solution to problems that are expected to behave more closely to the least absolute value solution as opposed to the least-squares solution. This sparse inversion is done through a framework called iteratively reweighed least-squares inversion (IRLS), where a set of model weights and data weights are used to guide the results to either the ℓ_1-norm or ℓ_2-norm solution. IRLS is implemented using a nested loop algorithm with the inner loop consisting of a standard conjugate gradient (CG) algorithm with the outer loops recalculating the weights for the model and data. The advantages ℓ_1-norm for the data space is that the inversion is less sensitive to sporadic events commonly associated with noise referred to as the robust solution. The ℓ_1-norm applied to the model space then pushes the model space to the solution with the most zero values favoring single large numbers instead of multiple small ones, this is commonly referred to as the sparse solution. Given a standard transformation operator that uses a mapping operator L that maps the data space d to the model space m (Claerbout, 1992)

$$\mathbf{d} = \mathbf{Lm}, \tag{4}$$

where to calculate the inverse we need to seek an m that fits the data while also minimizes

$$||\mathbf{m}||_{\mathbf{1}}^{\mathbf{1}}, \tag{5}$$

with the fit to the data determined by minimizing

$$||\mathbf{d} - \mathbf{Lm}||_{\mathbf{2}}^{\mathbf{2}}. \tag{6}$$

By adding model and data weights we can then dictate the algorithm to converge towards the least absolute value solution (ℓ_1) or least squares (ℓ_2) by reformulating the original least-squares equation through the minimization of the modified equation

$$\begin{gathered} ||\mathbf{W_m m}||_{\mathbf{2}}^{\mathbf{2}} \\ \text{Subject to } ||\mathbf{W_d}(\mathbf{d} - \mathbf{Lm})||_{\mathbf{2}}^{\mathbf{2}}. \end{gathered} \tag{7}$$

Where $\mathbf{W_d}$ is the data weights and $\mathbf{W_m}$ is the model weight. These weights allow the user to adjust the preference to either the residual or model, where $\mathbf{W_m}$ can be changed

to obtain a certain fit to either a sparse ℓ_1 or smooth ℓ_2 model, and $\mathbf{W_d}$ can be changed to prefer non-spontaneous events or "weight" all events equally in the input data. The weights for the matrices are generally diagonal with $\mathbf{W_d}$ constructed as

$$\begin{aligned} \text{diag}(\mathbf{W_d}) &= |\mathbf{r}|^{(p-2)/2} \\ \text{where } \mathbf{r} &= \mathbf{Lm} - \mathbf{d}. \end{aligned} \tag{8}$$

Where p is either 1 or 2 for inverting based on either the ℓ_1 or ℓ_2 data norm. $\mathbf{W_m}$ is normally constructed as

$$\text{diag}(\mathbf{W_m}) = |\mathbf{m}|^{(2-p)/2}. \tag{9}$$

Where p is either 1 or 2 for inverting based on either the ℓ_1 or ℓ_2 model norm. Due to the data weights being created through division a damping factor is chosen to avoid division by zero. The damping factor is determined as a percentile of the data.

For the inversion approach instead of using a single operator $\mathbf{L}$ we use two operators $\mathbf{\Gamma}$ and $\mathbf{R}$ as the blending operator and radon operator respectively. The relationship of which is dictated by

$$\mathbf{L} = \mathbf{\Gamma R}, \tag{10}$$

where the adjoint is represented by

$$\begin{aligned} \mathbf{L}^H &= (\mathbf{\Gamma R})^H, \\ \mathbf{L}^H &= \mathbf{R}^H \mathbf{\Gamma}^H. \end{aligned} \tag{11}$$

The main advantage of the inversion approach for deblending is that all of the signal and blending noise can be explained in other domains where they appear coherent, as opposed to the denoising approach which seeks to simply remove the blending noise frame by frame.

The denoising approach first applies pseudo-deblending to the blended data to return a data set in standard acquisition frames seen in Figure 1 b, which then allows for separation of the shots through denoising in the other domains. Due to the fact that blending noise is just interfering signal, the amplitude of the blending noise is the same as the signal's. The amplitude similarity results in an issue where the larger amplitude blended reflections would cover and destroy low amplitude events like those of diffractions or weak reflections in the data thus making it unrecoverable. This issue, however, does not affect the inversion approach as it seeks to remap all events back to their coherent frames, thus allowing for the recovery of low amplitude events. The general process to denoising using sparse/robust radon is to first apply the adjoint blending operator to deblend the data

$$\mathbf{D_{pseudo}} = \mathbf{D_{bl}} \mathbf{\Gamma}^H, \tag{12}$$

which then the pseudo deblended data is used in the inversion with the following objective function to denoise using the Radon transform

$$||\mathbf{D_{pseudo}} - \mathbf{Rm}||_p^p + \mu ||\mathbf{m}||_q^q, \tag{13}$$

Where $\mathbf{D_{pseudo}}$ is the pseudo deblended data, $\mathbf{R}$ is the Radon transform operator, $\mathbf{m}$ is the model, μ is the tradeoff parameter, and p and q are the data and model norms respectively. A graphical representation of radon denoising can be seen in Figure 5, using standard hyperbolic radon to down weight spontaneous events in the shot record in an attempt to recover to coherent signal.

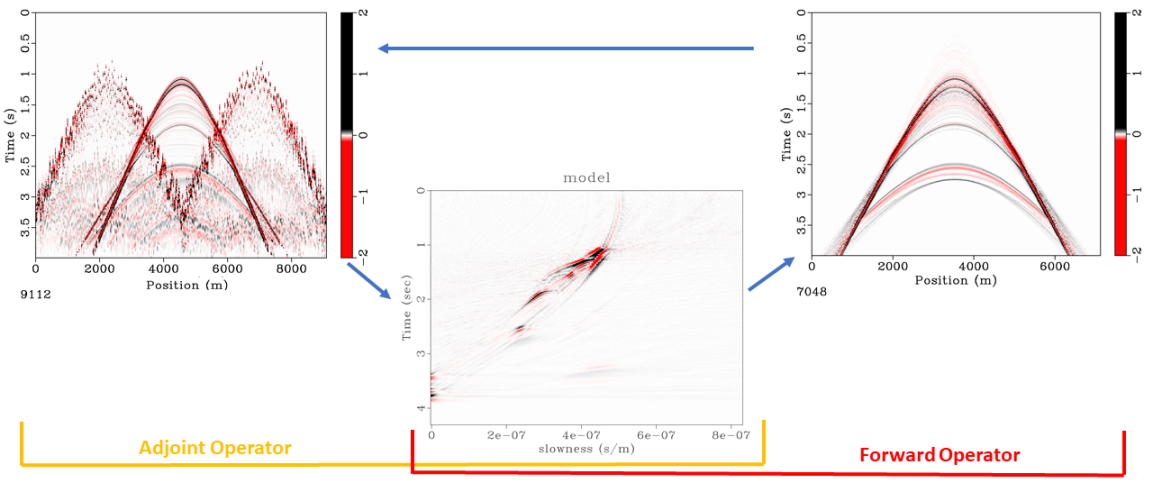

FIG. 5. Denoising approach to deblending: The denoising approach to deblending, using sparse/robust radon to remove the blending noise from each pseudo deblended reciever frame.

The inversion approach is seen in Figure 6, is not prone to issues like large-amplitude blended noise overpowering small-amplitude events, as it seeks to find a solution that fits all the data. Large blending noise, in this case, can be explained by refitting the blending noise back to their own data frames weather it is in CMP or receiver domain. With the amplitude explained, the lower amplitude events are then recovered. The only drawback to the inversion approach to deblending is the memory requirement, this is because the entire data set and the model must be held in memory to calculate the gradient on each iteration. In contrast, the denoising approach is much more lightweight due to only needing to hold the frame being denoised in memory rather than the entire data set. The inversion approach uses a different objective function as it seeks to incorporate both the blending operator as well as the Radon operator in the inversion, we can formulate the objective function given the generic equation

$$\begin{aligned} &||\mathbf{d} - \mathbf{Lm}||_p^p + \mu||\mathbf{m}||_q^q, \\ \text{where } &\mathbf{L} = \mathbf{\Gamma R}, \mathbf{d} = \mathbf{D_{bl}}, \\ \text{giving us } &||\mathbf{D_{bl}} - \mathbf{\Gamma R m}||_p^p + \mu||\mathbf{m}||_q^q. \end{aligned} \tag{14}$$

Where instead of using the pseudo deblended data, we input the blended data and have the operator contain both the blending and Radon operators.

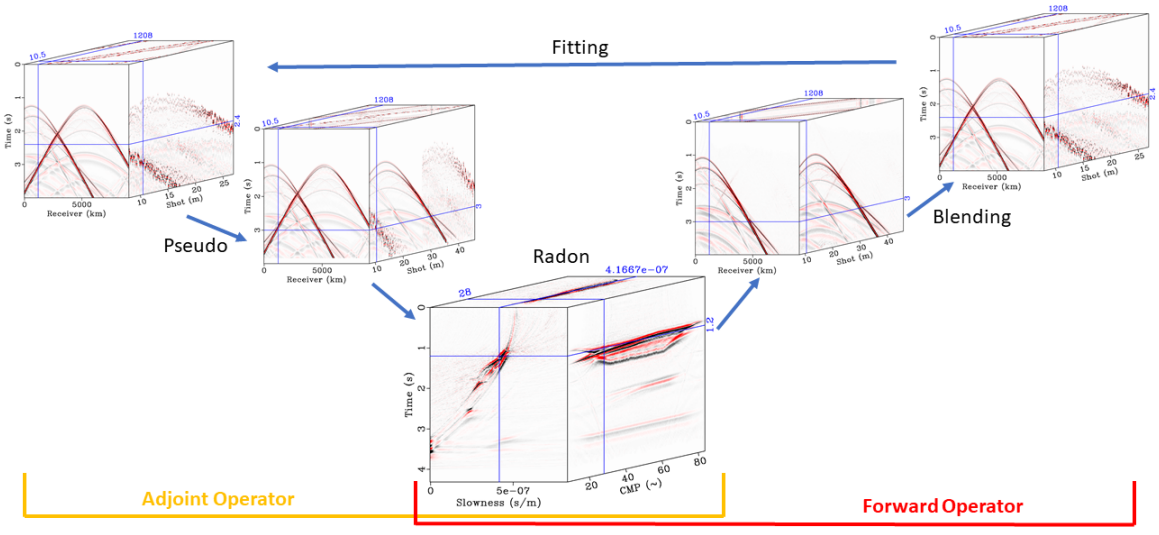

FIG. 6. Inversion based deblending: Using the inversion on both the radon and blending operators to remap data to their own frames, blending noise is mapped to coherent frames, all data is explained.

By implementing the deblending as an inversion based operation using the blending and Radon operators together, the system acts like an overcomplete dictionary, where the information is represented and redundant in the pseudo deblended frames. This aids in convergence as the redundant information explains each other through the remapping of data back to their respective shot frames.

EXAMPLES

To examine the performance of the inversion completed in the CMP domain, all inversion performed are using sparse constraints or ℓ_2-ℓ_1 for the data and model weights respectively. We created a set of increasingly complex finite difference data sets, blended both numerically and through finite-difference itself, the simplest of which is shown in Figure 8, which was performed on a two-layer wedge model with diffraction points planted along the model seen in Figure 7. The blending was done through finite-difference with two shots being fired at the same time with a randomized delay between 0 and 400 samples.

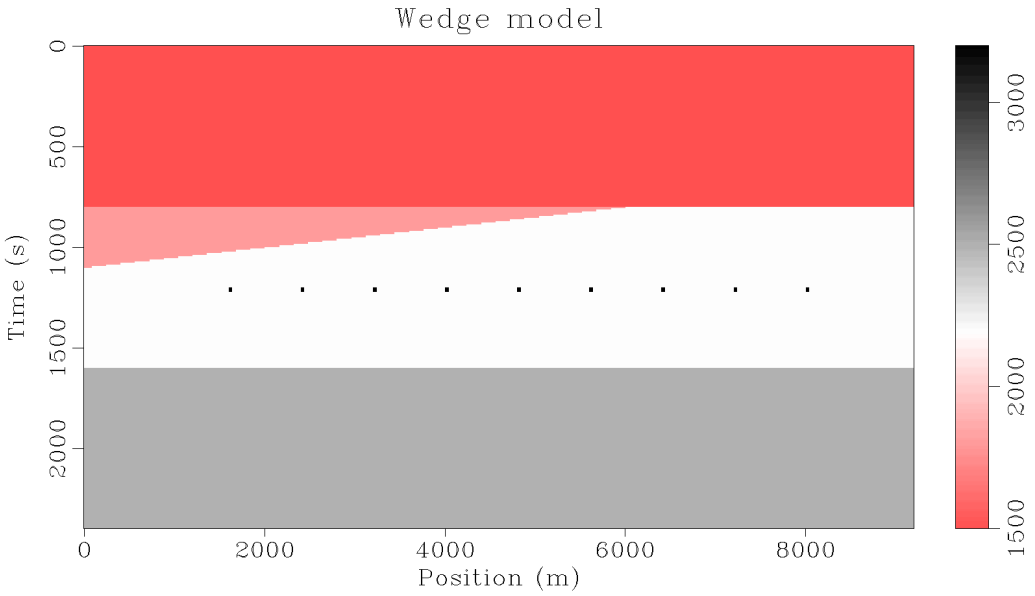

FIG. 7. Wedge model: Model consisting of two layers, a dipping refector(wedge), and a set of diffraction points.

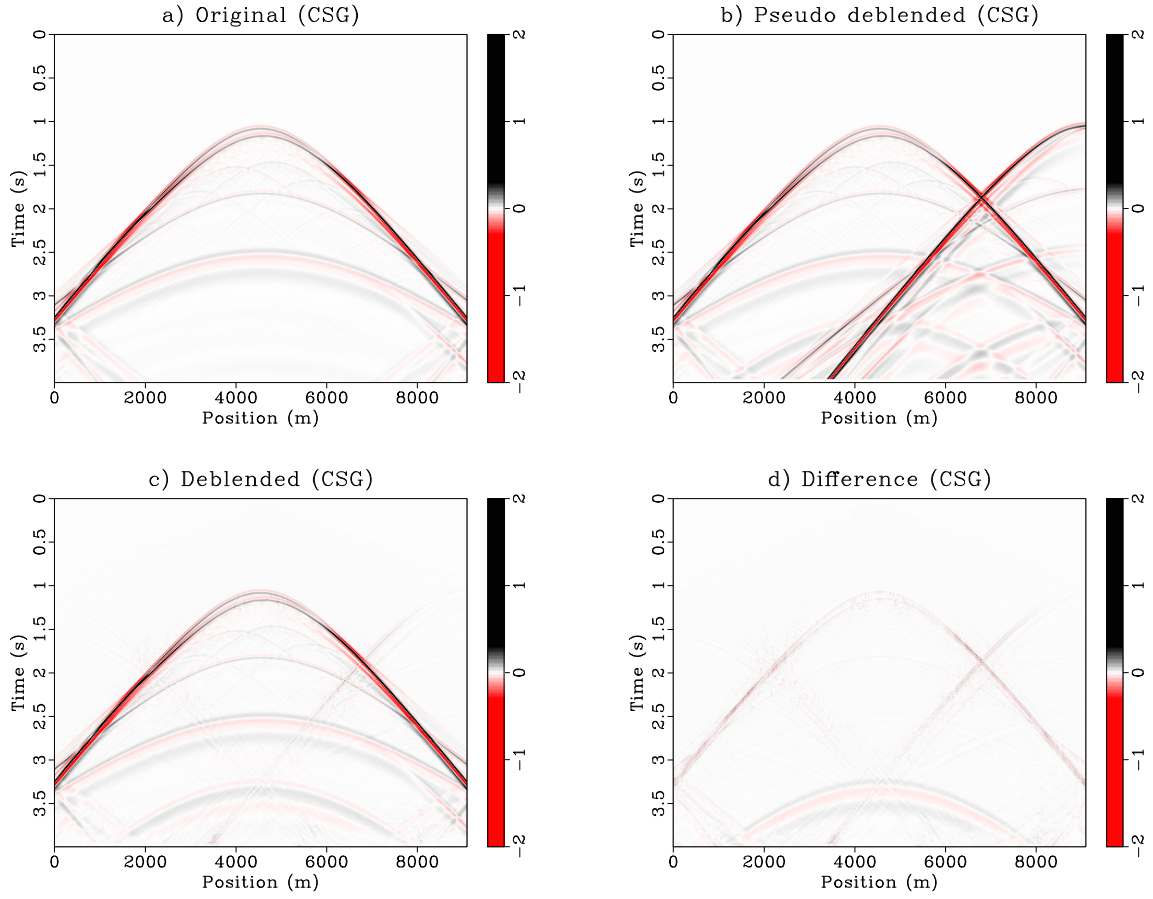

FIG. 8. Wedge model results: Deblending of a two simultanious source survey with a) unblended data, b) pseudo-deblended data, c) deblended data, and d) differnce respectively.

The CMP domain of the pseudo deblended data can be seen in Figure 9b where the blending noise can be seen overlapping the signal at the sides and later arrival times. The results of the deblending in shot domain can be seen in Figure 8c and CMP domain in Figure 9c. With the pseudo deblended original in Figure 8b.

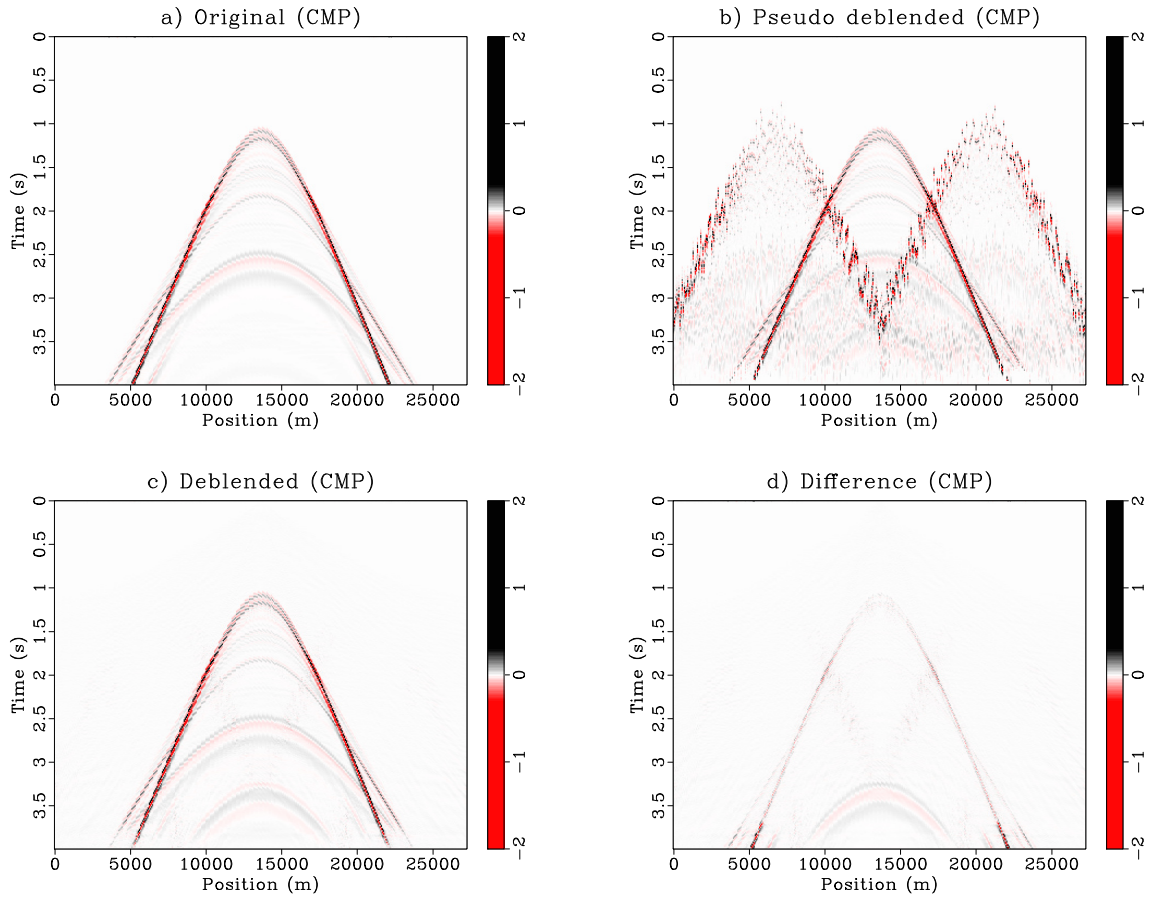

FIG. 9. Wedge model results in CMP: Deblending of a two simultanious source survey in CMP domain with a) unblended data, b) pseudo-deblended data, c) deblended data, and d) differnce respectively.

While there is still some blended signal remaining in the blended shot, the deblending preserved the in-focus shot well, including the diffractions which can be seen near the top of both Figures 8b and 8c. The inability to remove the entire influence of the out of focus shots may be due to over tuning of the sparsity of the Radon transform.

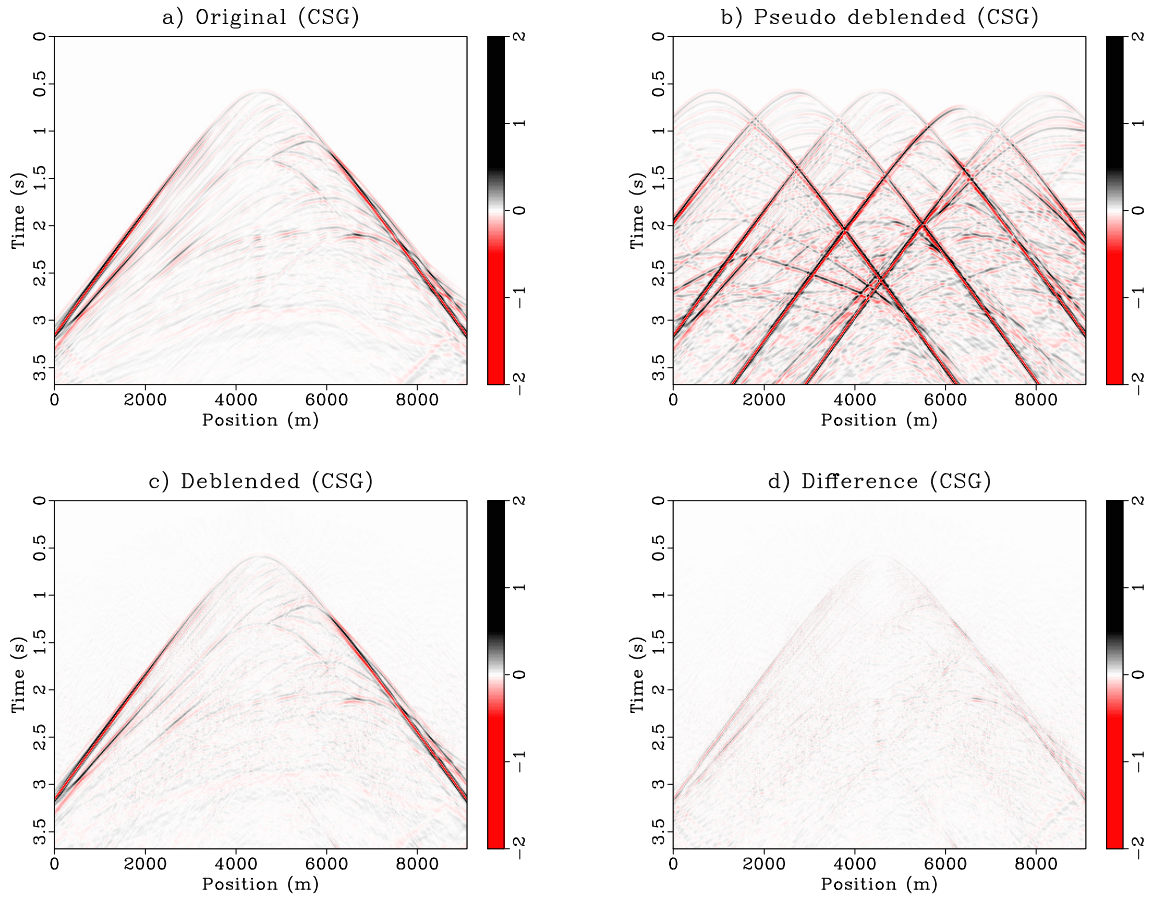

FIG. 10. Marmousi results: Results of the marmousi model inversion with 5 simultanious shots, with a) unblended, b) blended, c) deblended and d) difference respectivly.

The next example tested was the marmousi model (Brougois et al., 1990), a more difficult blending schedule was used to test the limits of deblending in the CMP domain. For the marmousi examples, 5 shots were simultaneously fired instead of two and firing delay was randomized between 0-200 samples with an example shot before and after deblending provided in Figure 10. The increased difficulty in deblending the shots can best be seen in the CMP domain in Figure 11c where most of the signal is covered with blending noise. In this condition, deblending by denoising would most likely fail as a significant amount of the reflections are overpowered by blending noise, in this situation inversion would be the most effective solution.

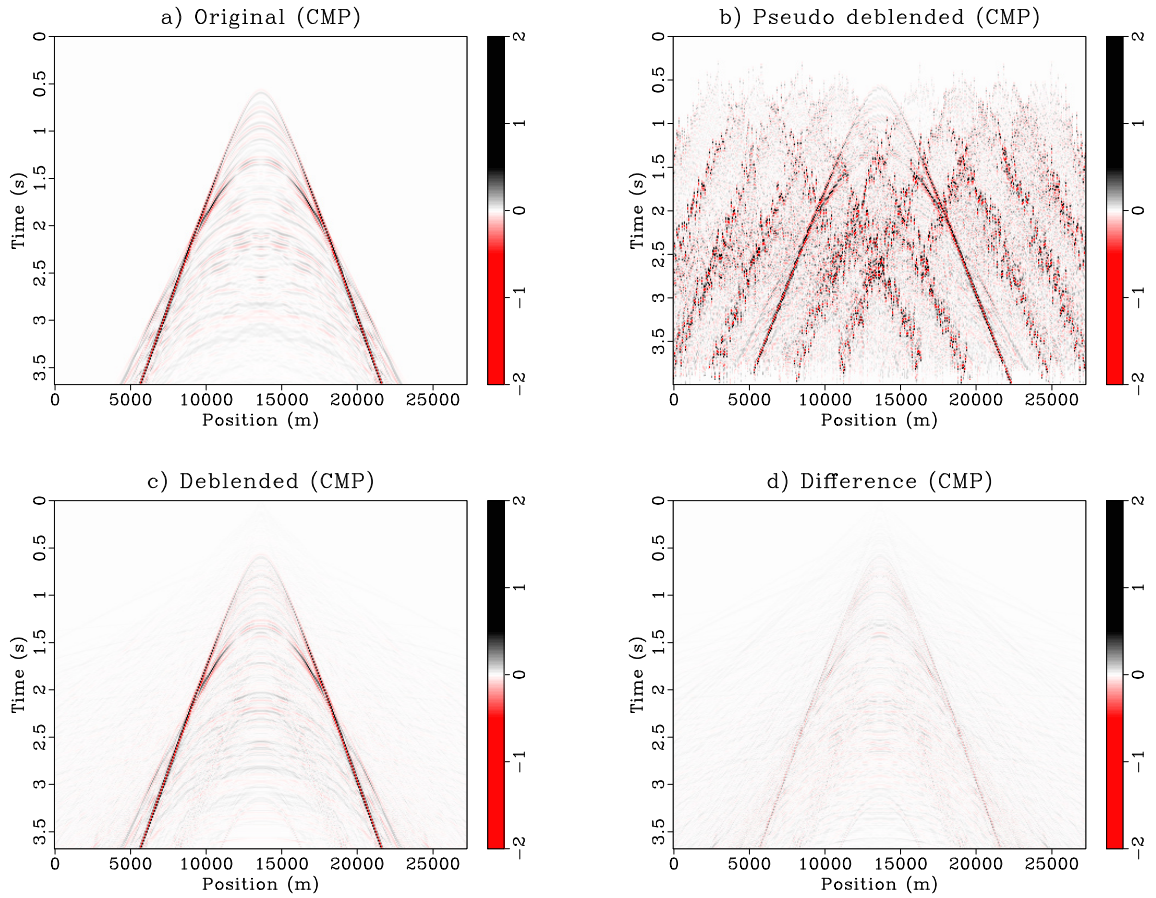

FIG. 11. Marmousi results CMP: Results of the marmousi model inversion in CMP domain with 5 simultanious shots, with a) unblended, b) blended, c) deblended and d) difference respectivly.

The results of the inversion of the marmousi model show that lower amplitude reflections are capable of being recovered while completely buried by blending noise from the 4 interfering shots, seen in Figure 11 with the primary issue being that not all blending noise is removed, this is further corroborated in Figure 10c where most of the events are recovered with only some blending noise remaining.

The results for the marmousi model show close to a worst-case scenario for testing, with a high number of simultaneous shots while having small delay times, the only notable issues with the results are not missing events but event amplitudes, we believe that this is either due to enforcing too much sparsity in the inversion or that the amplitude variation across an event is too much for the basis function to properly map.

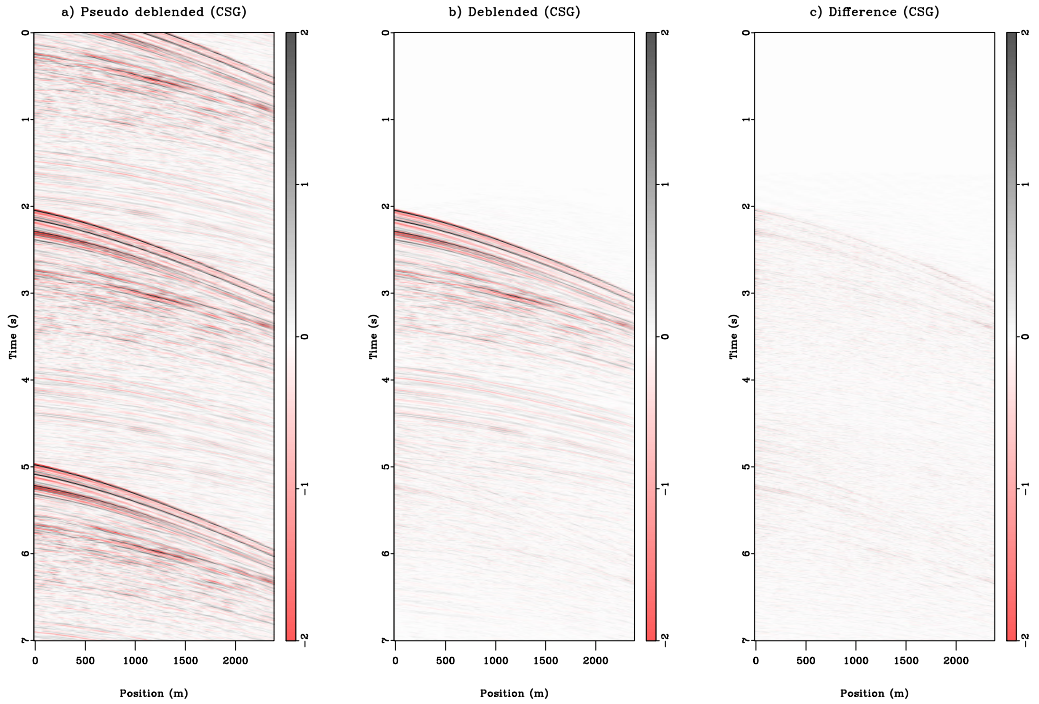

FIG. 12. Gulf of Mexico data: Deblending using a gulf of Mexico marine dataset with an shot time interference of 70%, with Blended data in a), deblended in b), and difference in c).

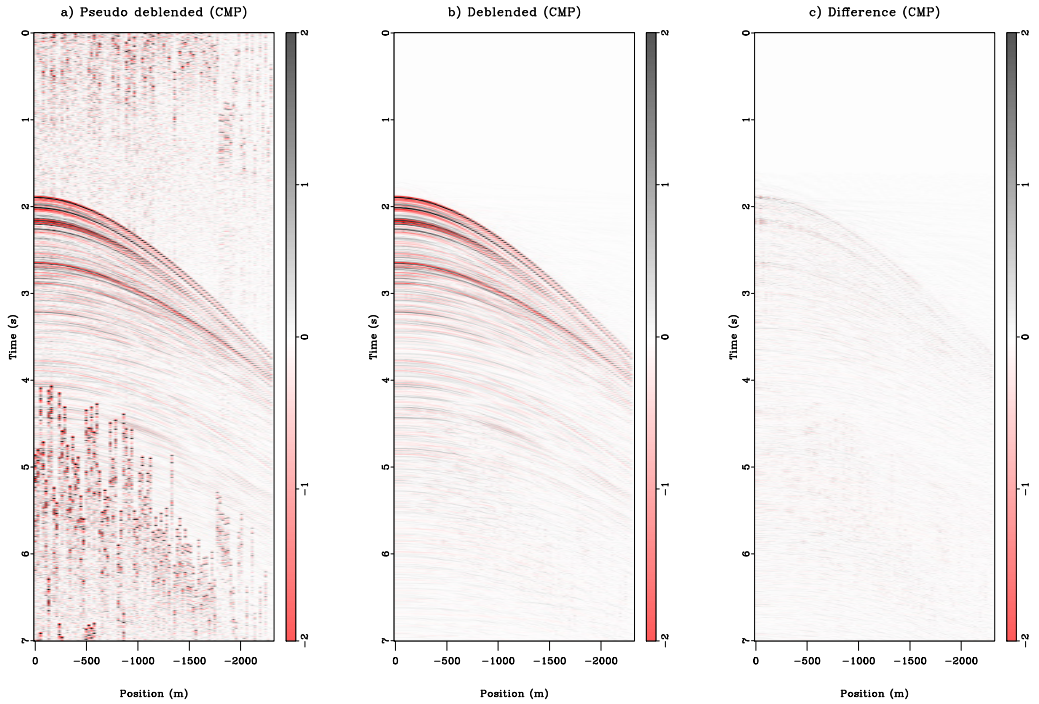

FIG. 13. Gulf of Mexico data in CMP domain: Deblending using a gulf of Mexico marine dataset with results in the CMP domain, with Blended data in a), deblended in b), and difference in c).

The final example tested was a gulf of Mexico dataset (Ibrahim, 2015), this real-world dataset was blended numerically with a firing time overlap of 70% using continuous lis-

tening with 90 receivers and 90 shots. The results of the deblending of the gulf of Mexico dataset can be seen in with difference plots in Figure 12 and CMP domain in Figure 13. An enlarged set of results can be seen in Figure 12b and in the CMP domain in Figure 13b. Looking at the Gulf of Mexico results, the deblending is very effective at separating each shot from one another. Recovery of events is also quite good with the only issue being that the first arrivals of later shots are not completely removed seen in the shot domain at around 5.2 seconds where some amplitude is left behind and not properly separated.

CONCLUSION

The proposed inversion approach of deblending in the CMP domain allows for the separation of blended data by explaining all the data and remapping blended shot back to their own domain. This allows us to recover low amplitude events that are obscured by blended noise. By implementing a sparse constraint to the inversion algorithm, we can converge at a unique solution that maximizes the coherency of the signal in hyperbolic arrivals. The use of the CMP domain allows us to implement a simple hyperbolic Radon operator. This allows us to reduce the computational time of the inversion while allowing us to stay in the time domain avoiding aliasing concerns associated with frequency-domain methods. Therefore the inversion approach in the CMP domain is a viable way to deblend simultaneously sourced signals while recovering low amplitude events.

ACKNOWLEDGMENTS

We thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors, and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 461179-13.

REFERENCES

Abma, R., Howe, D., Foster, M., Ahmed, I., Tanis, M., Zhang, Q., Arogunmati, A., and Alexander, G., 2015, Independent simultaneous source acquisition and processing: GEOPHYSICS, **80**, No. 6, WD37–WD44.

Beasley, C. J., Chambers, R. E., and Jiang, Z., 1998, A new look at simultaneous sources, 133–135.

Berkhout, A. J., 2008, Changing the mindset in seismic data acquisition: The Leading Edge, **27**, No. 7, 924–938.

Berkhout, A. J., Blacquière, G., and Verschuur, D. J., 2009, The concept of double blending: Combining incoherent shooting with incoherent sensing: GEOPHYSICS, **74**, No. 4, A59–A62.

Brougois, A., Bourget, M., Lailly, P., Poulet, M., Ricarte, P., and Versteeg, R., 1990, Marmousi, model and data, *in* EAEG workshop-practical aspects of seismic data inversion.

Claerbout, J. F., 1985, Imaging the Earth's Interior: Blackwell Scientific Publications, Ltd., Oxford, UK, UK.

Claerbout, J. F., 1992, Earth soundings analysis: Processing versus inversion, vol. 6: Blackwell Scientific Publications London.

Garottu, R., 1983, Simultaneous recording of several vibroseis seismic lines, 308–310.

Ibrahim, A., 2015, Separating simultaneous seismic sources using robust inversion of Radon and migration operators: PhD Thesis, University of Alberta.

Ibrahim, A., and Sacchi, M. D., 2014, Simultaneous source separation using a robust Radon transform: Geophysics, **79**, No. 1, V1–V11.

Ibrahim, A., and Sacchi, M. D., 2015, Fast simultaneous seismic source separation using stolt migration and demigration operators: GEOPHYSICS, **80**, No. 6, WD27–WD36.

Mahdad, A., Doulgeris, P., and Blacquiere, G., 2011, Separation of blended data by iterative estimation and subtraction of blending interference noise: GEOPHYSICS, **76**, No. 3, Q9–Q17.

Stanton, A., and Wilkinson, K., 2018, Robust deblending of simultaneous source seismic data.

Thorson, J. R., and Claerbout, J. F., 1985, Velocity-stack and slant-stack stochastic inversion: GEOPHYSICS, **50**, No. 12, 2727–2741.

Trad, D., Hargreaves, N., verWest, B., and Wombell, R., 2004, Multiple attenuation using an apex shift radon transform: SEG Expanded Abstracts.

Trad, D., Siliqi, R., Poole, G., and Boelle, J.-L., 2012, Fast and robust deblending using Apex Shifted Radon transform, 1–5.

Trad, D., Ulrych, T., and Sacchi, M., 2003, Latest views of the sparse radon transform: GEOPHYSICS, **68**, No. 1, 386–399.

Trad, D. O., 2015, Least squares kirchhoff depth migration: implementation, challenges, and opportunities: SEG Technical Program Expanded Abstracts 2015, 4238–4242.

Urruticoechea, C. R., 2015, Seismic blending and deblending of crossline sources: M.Sc. thesis, Delft University of Technology.