A Madagascar package for deblending in multiple flavours

Daniel Trad

ABSTRACT

Data acquisition is the most expensive part of seismic methods and the need to decrease its costs leads to cost-effective compromises. In particular, 3-dimensional data surveys always have deficient sampling in at least 2 of the 4 spatial dimensions. Offset and azimuth dimensions are always undersampled and migration algorithms produce suboptimal angle and offset gathers, which affects AVO and AVAz studies. As a consequence, geophysicists make extensive efforts in the mitigation of these sampling problems. These efforts usually involve two directions: data interpolation and simultaneous acquisition. Interpolation is intended to create new seismic traces from the acquired samples by using sparse transformations. Simultaneous acquisition, also known as blending, attempts to mitigate the sampling problem by acquiring more data in less time by recording several shots simultaneously. Its main difficulty is the processing of the resulting seismic data, which requires either shot separation very early in the signal processing chain or the development of new processing/inversion capabilities that can work directly with blended data.

There are several approaches for deblending seismic data: denoising, inversion and physical transforms. From a general point of view, the three are closely related but they differ in how and which transformations are used. In this paper, I will discuss the relationships between these approaches from an algorithmic point of view.

INTRODUCTION

Seismic processing has a long history of compromises between best science and practical options. Examples are 2D vs 3D acquisition, postack vs prestack migrations, time vs depth migrations, acoustic vs elastic data, interpolation vs denser acquisition. Many successful seismic technologies appeared as compromised solutions where economical or computational savings justified relatively poor scientific choices.

Often, technological advances make some of these approximations unnecessary. Sometimes, these approximations keep their merits for new reasons. An example is time migration, originally invented as a manner to reduce computational cost, but still very useful as a way to increase robustness to velocity model inaccuracies. In other cases, computational advances have brought us closer to most advanced technologies but still not enough to make them practical for every day uses, like elastic migration and full waveform inversion. These technologies are refinements on our toolboxes and are usually applied as a final step of the divide and conquer strategy that has been successful in seismic for a century.

New technologies have become essential when more complex economical or social factors come into play. A few years ago, seismic was evolving clearly in the direction of denser acquisitions, with longer offsets, wider and denser azimuths, wider frequency ranges and denser shots and receivers. As economical and social factors came in, the effort shifted toward cheaper and less damaging acquisitions. Breakthroughs in seismic have become more targeted towards achieving the same quality with less cost and damage, rather than just more quality.

Blended acquisitions are an example of this tendency. Shot/receiver density is an unequivocal propeller for higher section quality and interpretation accuracy. Blended acquisitions promise to increase shot/receiver density without additional cost. How realistic is that promise? It quite depends on the level of enthusiasm of the practitioners. As seismic processing/inversion researchers, we can only do the best we can to overcome the challenges, which leads to the main topic of this report.

The critical challenge of deblending technology is one of two options:

1. how to separate blended data such that can be processed with standard techniques. This involves separating raw data based on coherence measures, which are often compromised until later on the processing sequence.
2. how to process a blended data set as is. This involves redesigning our processing/inversion toolboxes so they can work directly on blended data.

Most practical applications so far have been of the first type. Even if researchers can successfully implement the second approach, most likely we will see industrial applications lagging in the actual use of such technology because of the heavy load of legacy tools and other issues including confidence.

However, it is possible to combine both approaches efficiently. For example, we can use the first approach to extract information about velocities, statics, and other corrections, and resort back to the original blended data to apply migration directly. This is quite possible considering that migration of blended data is already in practical use as a way to reduce computational cost. We will see some examples at the end of this report. Regarding the separation itself (1), we can think of three related but distinct approaches, all of which map the blended data into an extended (separated) space:

- Denoising: data are decomposed on groups, each one showing only one of the shots as coherent energy and the others as incoherent. The mechanics of which shots appear coherent and which ones not, depends on the particular blending and transform methodology.
- Inversion: data are decomposed in groups, similar to before, but all groups are simultaneously inverted to a transform space with a suitable operator. These groups are all fitted (predicted) by the combination of transform panels. This is a true parallel transform design, where shots are predicted simultaneously.
- Physical transforms: use physics to map the blended data to the correct physical space to which they belong. Once the correct physical space is built, shots can be predicted in the deblended (extended) space.

IMPLEMENTATION AND GENERAL DIRECTIONS

Deblending package

In this report, we will see how the three approaches are related by looking at how they are implemented. We will see that converting between different techniques can be explained in terms of either redefining the transform or just moving the transform from outside to the inside of the optimization algorithm. The main geophysical framework used for the implemented package is Claerbout (1992). All programs used in this report are written in C++ using an Object-Oriented Programming approach. I used the "Madagascar" framework for convenience and testing but code can be compiled independently with any C++ (std=+11) compiler by just replacing I/O functions and the parameter interface. I use C++ and classes to encapsulate ideas into logical units. Each module contains similar components, each of which is a class: a driver, a blender, a solver and a transform. Differences between approaches reduce to switching classes in a module. Replacing the transform class should be relatively easy since all the other components should not change. An improvement on this package would be to use inheritance to minimize code duplication.

To deblend or not to deblend?

The big promise from blended technology is it allows us to increase illumination without increasing acquisition costs. However, processing blended data is challenging. Each seismic trace contains more than one shot, which means also more than one offset, one azimuth and one CMP. In Figure 1 we see a two-vibroseis blended data acquired by CREWES in September 2018. This is a challenging data set to process because there is no distinction between the energy coming from the two shots. We will see later an example of deblending these data using a Radon transform.

As explained by Berkhout (2008); Berkhout et al. (2009); Verschuur (2011), at the highest level, there are two different possible approaches when facing blended data sets. As illustrated in Figure 2 (left), we can apply a separation procedure (de-blending) to the input data, and then process the deblended data in the usual way. This is the most common approach taken in the industry. However, we can also work directly with the blended data after re-defining the signal processing modules required for processing (Figure 2, right). This is more difficult because implies a significant redesign of the software. To my knowledge, most of these processes have not been developed yet, except for blended migration (marked in red in Figure 2). We will see later that blended migration works well provided that the velocity model is accurate. That suggests that a key challenge is to develop blended data velocity analysis tools. But this will also require other components of the seismic processing chain, for example, blended-multiple removal and in the case of land data, blended-statics tools. A third approach is possible: deblend data as in the left workflow, use these data to get the necessary information, and then return to the blended data before migration. The advantage of this hybrid approach is that it permits some deficiencies during the deblended procedure, as long as the information obtained from these deblended temporary data is not seriously compromised.

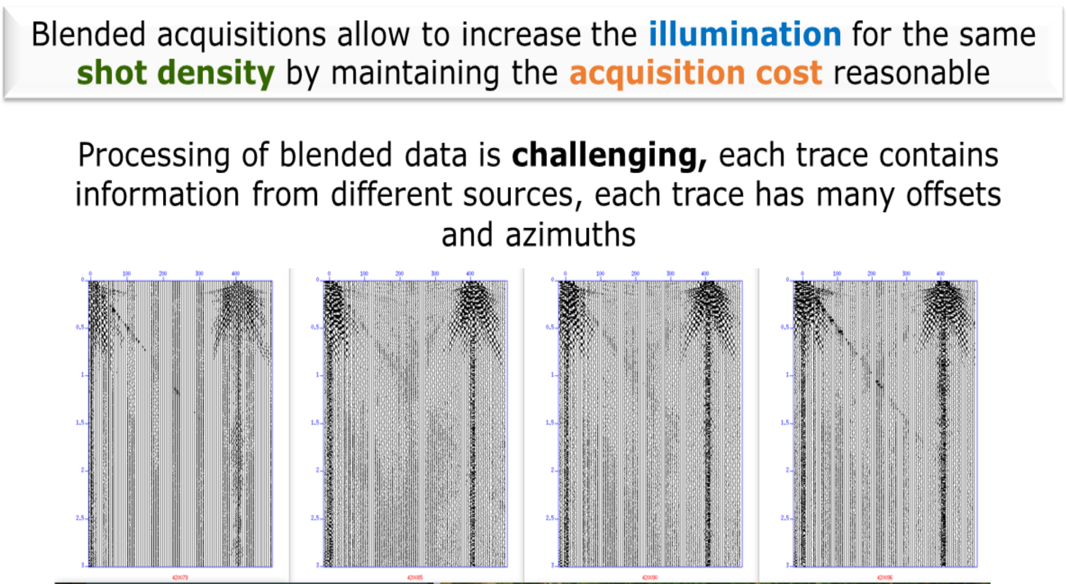

FIG. 1. Deblending challenges. CREWES data acquired in 2018

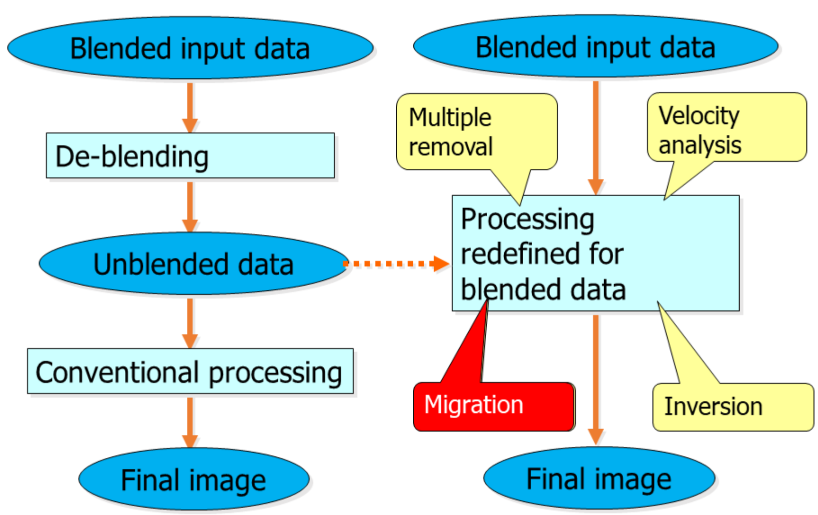

FIG. 2. Two different approaches for blended data processing (Verschuur, 2011).

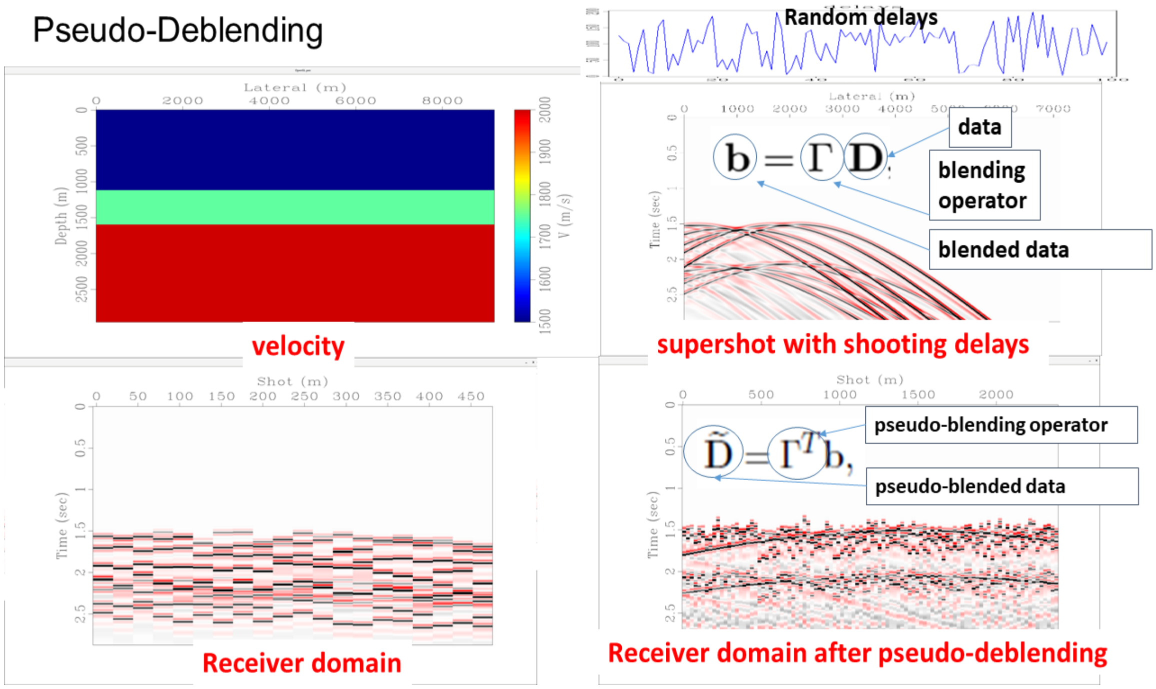

FIG. 3. When applying the pseudo-deblending operator, coherence is restored only for the particular shot whose delay is correctly compensated.

DEBLENDING APPROACHES

Simultaneous shooting greatly reduces the cost of information (Beasley et al., 1998) but it brings significant difficulties because most processing algorithms use a unique source location, either by working on shot gathers or other groups derived from acquisition coordinates (for example offset and common midpoint gathers, inlines, crosslines). Therefore, it is necessary to separate at front the information coming from different shots, which is the process known as "deblending". Although there are many deblending tools available, often they use data properties that are not well defined before processing. For example, methods based on coherence and sparseness may not work in raw data before static corrections are applied. Although the principles behind many of these deblending tools have strong similarities with regularization and denoising techniques, the challenge comes from the need to perform this operation on raw data at the very beginning of the processing sequence.

Deblending by denoising and inversion

There are many different deblending techniques. Good introductions can be read in the following papers by the deblending pioneers: Beasley (2008); Abma et al. (2015); Berkhout et al. (2009). The first step to understand deblending techniques is to consider "pseudo-deblending" (Mahdad et al., 2011), which is the simplest possible approach and a fundamental component of other techniques. Figure 3 shows a simple velocity model (top-left), and a blended super-shot (top-right). This super-shot is formed by several simultaneous shots (five in this case). A typical data set will contain many super-shots. Each super-shot can be thought of a mixing of several individual shots:

$$\mathrm{b} = \Gamma \mathrm{D}. \tag{1}$$

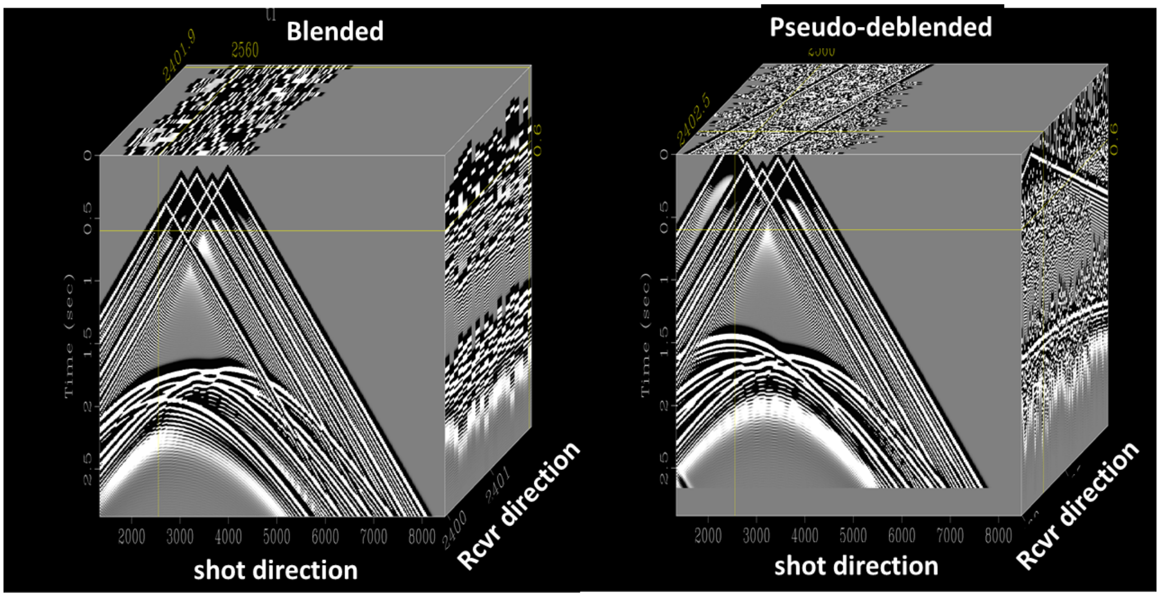

FIG. 4. Volume expansion produced by pseudo-deblending.

In this equation, $\mathbf{b}$ is a super-shot, $\mathbf{D}$ contains several individual (regular) shots and $\mathbf{\Gamma}$ is a mixing-delay matrix. Typically $\mathbf{D}$ is larger than $\mathbf{b}$, and therefore $\mathbf{\Gamma}$ would have more columns than rows (5 times more in this case). Usually, this mixing matrix will have a time delay operation applied to each column. The delays, shown in Figure 3 on the top right, would normally be random. As a consequence of these random delays seismic events will look incoherent in any other domain but the shot domain. For example, on the bottom left we see the blended data sorted in receiver gathers. To recover the coherence in this domain, we need to remove the delays. But each trace in a receiver gather contains many shots, five in this case, and each shot has a different delay, so which of the five delays should we choose? The answer is all of them, one at a time, and each choice will produce a trace corresponding to a different possible survey.

If the above paragraph reads like a riddle, a cleaner explanation can be obtained by using the forward-adjoint operator pair formulation (Berkhout et al., 2009; Mahdad et al., 2011; Urruticoechea, 2015). Since $\mathbf{\Gamma}$ can be thought of as a matrix, its adjoint can be expressed as $\mathbf{\Gamma^H}$ (transpose complex conjugated matrix), and a kinematic approximation of $\mathbf{D}$ can be quickly calculated as:

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

$\tilde{\mathbf{D}}$ are the pseudo-deblended data, which contains the right dimensions but not quite the correct separation. The effect of the transpose and conjugation will be to exactly remove the delays applied during mixing for one of the shots in each supershot. The other shots will appear at the wrong locations (incorrect delays). The effect in the original shot domain is not easy to see, but in the receiver or CMP domain it is quite clear. One of the many (five) possible surveys become coherent (Figure 3, bottom right). Figure 4 shows the resulting volume expansion and coherence gain obtained by apply $\mathbf{\Gamma^H}$ to the blended data. The data dimension has exploded by the number of blended shots, five in this case.

Although pseudo-deblending is not quite what we want, it puts the data in a format that

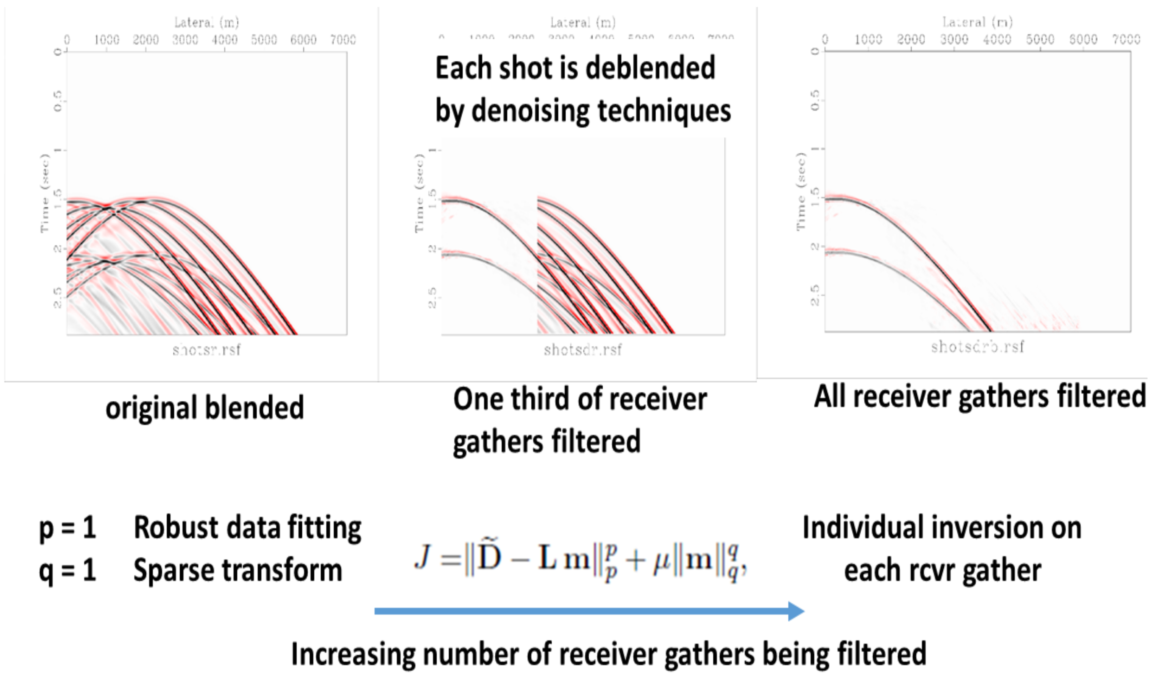

FIG. 5. Deblending by denoising techniques

we can "almost" handle with traditional denoising techniques. Now we just need to separate the coherent from the incoherent shots. For example, we can use any denoising technique that favours coherence, like Radon transform, Fourier transform, prediction filters, etc:

$$J = \|\mathbf{\Gamma^H b} - \mathbf{Lm}\|^1 + \lambda\|\mathbf{m}\|^1 \tag{3}$$

In this equation, J is a cost function that we will minimize, $\mathbf{L}$ is the particular chosen transform and $\mathbf{m}$ is a particular model resulting from this transform. As usual, the first part of J contains the data fitting term and the second part is the regularization term, weighted by a scalar coefficient λ that controls the regularization trade-off (more fitting less sparse or vice-versa) . The energy of each term is calculated by a ℓ_1 norm. In the residual term, the ℓ_1 norm tolerates outliers or large isolated residuals, which are in this case the energy coming from blended shots (Ibrahim and Sacchi, 2013). In the regularization term, the ℓ_1 favours transforms with a few large coefficients (sparse). Notice that minimization of this cost function requires an optimization algorithm like conjugate gradients (CG) to invert for the model $\mathbf{m}$. In this problem, the CG algorithm will operate independently on each shot in $\tilde{\mathbf{D}}$.

Figure 5 shows an example of this implementation when $\mathbf{L}$ is a hyperbolic Radon transform in the time domain (Trad et al., 2002). The figure on the left illustrates one of the many super-shots, the figure on the right is one of the separated shots, when all receiver gathers have been filtered. The figure on the centre illustrates one shot in the "middle" of the process when one-third of the receiver gathers have been filtered.

Although in principle this procedure works reasonably well for simple models, like the one in Figure 3, it tends to fail for more complex structures. One reason is that the random noise produced by the blended energy is high amplitude, the same as the signal. Therefore, this method is equivalent to denoising data with a signal to noise ratio equal to one. To get around this problem, we need to think of the blended energy not as incoherent unexplained data but as signal explained by the modelling operator (Abma and Yan, 2009). What kind of operator can explain random noise? None. However, the random noise is not random when the proper delays are applied. So really what is needed is a "parallel" data fitting, where several operators fit the data simultaneously, each operator with the correct delays for one of the five simultaneous surveys. Although it seems complicated, the implementation of such a parallel transform is not too different from the denoising algorithm (Equation 3). The new cost function we need to minimize is:

$$J = \|\mathbf{b} - \mathbf{\Gamma L m}\|^2 + \lambda \|\mathbf{m}\|^1 \tag{4}$$

There are only three differences between Equation 4 and Equation 3:

- The Γ operator is now part of the modelling operator.
- The input data are not the pseudo-deblended gathers $\tilde{\mathbf{D}}$ but the blended data $\mathbf{b}$.
- The residual term does not need an ℓ_1 norm but the standard ℓ_2 applied for regular Gaussian distributions.

What are the consequences of these three changes? The iterative algorithm that before was used independently on each gather to calculate the least squares transform, now has to be applied to all the gathers simultaneously. Because now we are fitting the data in "parallel" with many transforms simultaneously, we need to solve the inversion for all gathers at the same time. The previous "one-gather-at-time" inversion used to minimizing the cost function in Equation 3 becomes a all-at-once inversion where gradients and step sizes depend on all gathers simultaneously. The deblending of the whole survey is now a global problem with one minimum. This is similar to how we solve Least-Squares Migration, and we will come back to this similarity later. Since the outliers in the residuals are not unexplained any longer, we can drop the ℓ_1 minimization and use standard ℓ_2.

Figure 6 shows the equivalent result for Figure 5 but using this time inversion with the full-blown parallel operator. Figure 7 shows a comparison of the two methods. These two results are very similar because this data set comes from a very simple structure. As we will see next, results are quite different when data are complex.

As mentioned earlier, receiver gathers show the incoherent nature of blended energy when traces have random delays. CMP gathers also show the same effect, but usually, their sampling is coarser than for receiver gathers. However, CMP gathers have an advantage over receiver gathers for dipping or curved layers or diffractions: hyperbolic events tend to have their apexes near zero offset in CMPs. Non-zero apexes are problematic for any algorithm that uses offsets, like Hyperbolic Radon transforms or NMO based transforms. An apex shifted Radon transform (ASRT) (Trad, 2003) helps to address this. In practice, for deblending with hyperbolic Radon transform engines, anything more complex than flat layers require deblending in CMP gathers or ASRT (Trad et al., 2012; Ibrahim and Sacchi, 2015).

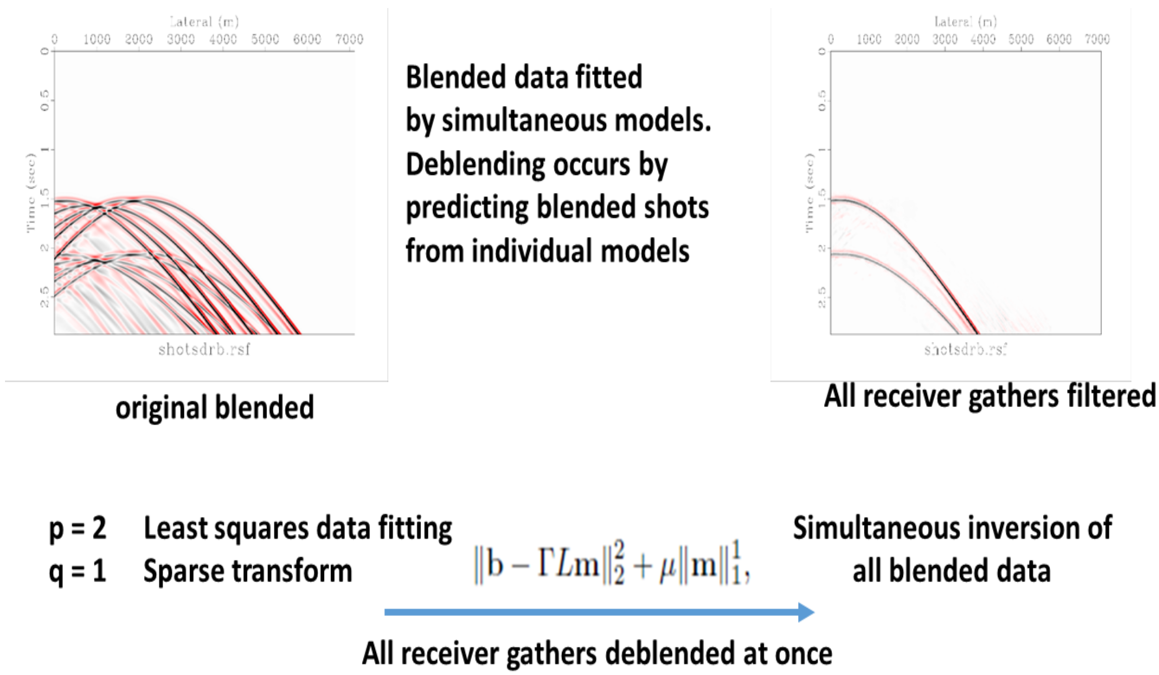

FIG. 6. Deblending by inversion techniques

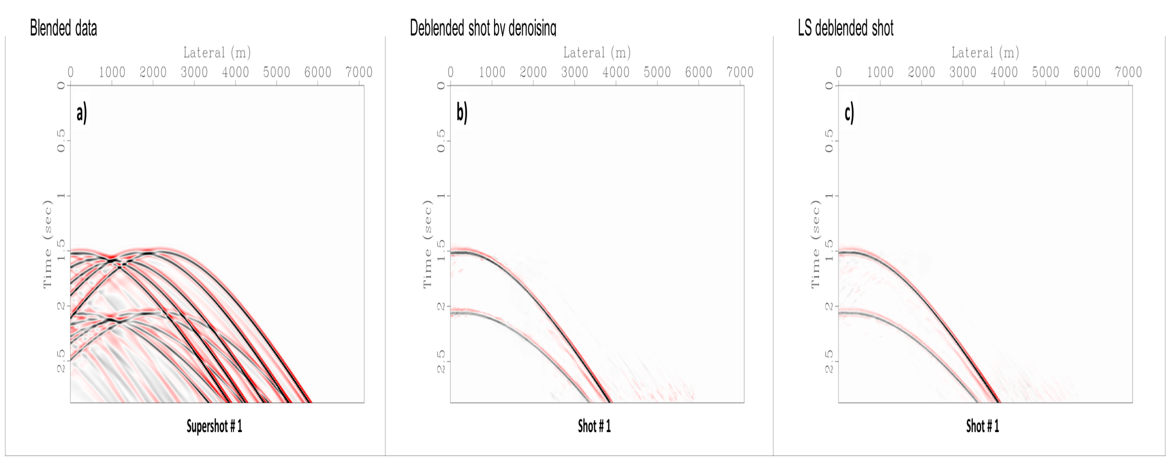

FIG. 7. Deblending by denoising and inversion techniques: a) One supershot (5 blended shots), b) One shot separated by denoising, c) The same shot separated by inversion.

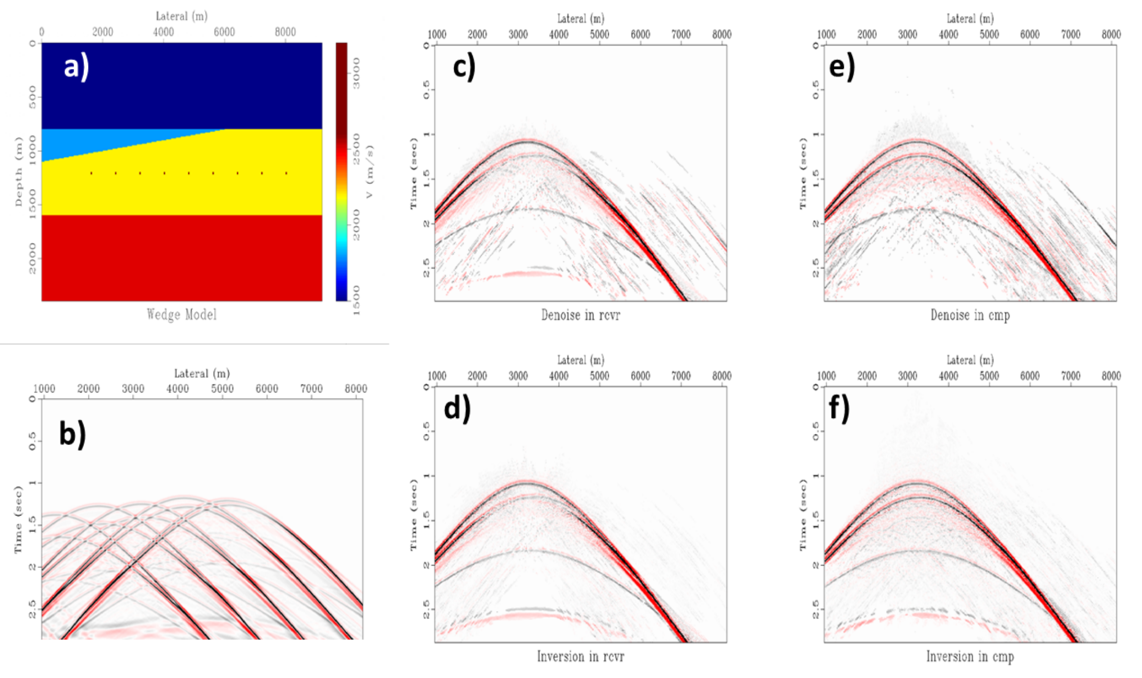

FIG. 8. Pushing the algorithms to the break-point: Deblending by denoising and inversion in CMPs and Receiver gathers for diffractions under a wedge.

We are not sure yet whether the standard hyperbolic Radon transform is flexible enough for complex structures. Let us compare results in CMPs and receiver gathers for both inversion and denoising. Let us push these methods to the break-point in difficult cases. Figure 8a shows a wedge model with scatterers underneath the dipping later. Figure 8b shows one of many simulated supershots, each one with five shots. Although still a simple model, it creates apex shifted events and diffractions with high curvature. Figure 8c and d show denoising and inversion in receiver gathers, Figure 8e and f show the same in CMP gathers. Inversion in CMP gathers is the most accurate and seems to reproduce well the dipping layer reflections, but diffractions seem attenuated. This is a challenging case for which our programs are still struggling.

Let us push the algorithms even more with a realistic salt example. Figure 9a shows a complex salt structure and Figure 9b shows one of 20 supershots, each one with 5 shots, from the left part of the model. Figure 9c shows for comparison a never-blended shot. Because of the large apex shift in the seismic events due to the complex structure, only the CMP domain can possible work with the hyperbolic Radon transform. Even in the CMP domain, the complexity of the events can overwhelm the prediction power of the Radon transform. Figure 9d shows the predicted blended data from least squares (b in Equation 4). Figure 9e shows the predicted shot. Most blended energy has been separated but the most complex parts are not yet completely separated. Figure 9f shows one CMP gather before the deblending. Events are heavily entangled and its Radon transform Figure 9g shows significant leakage of energy. Enforcing more sparseness may clean it better but also lead to some missing weak energy. That is the trade-off that deblending algorithms have to struggle with. We haven't yet found a fit-all solution.

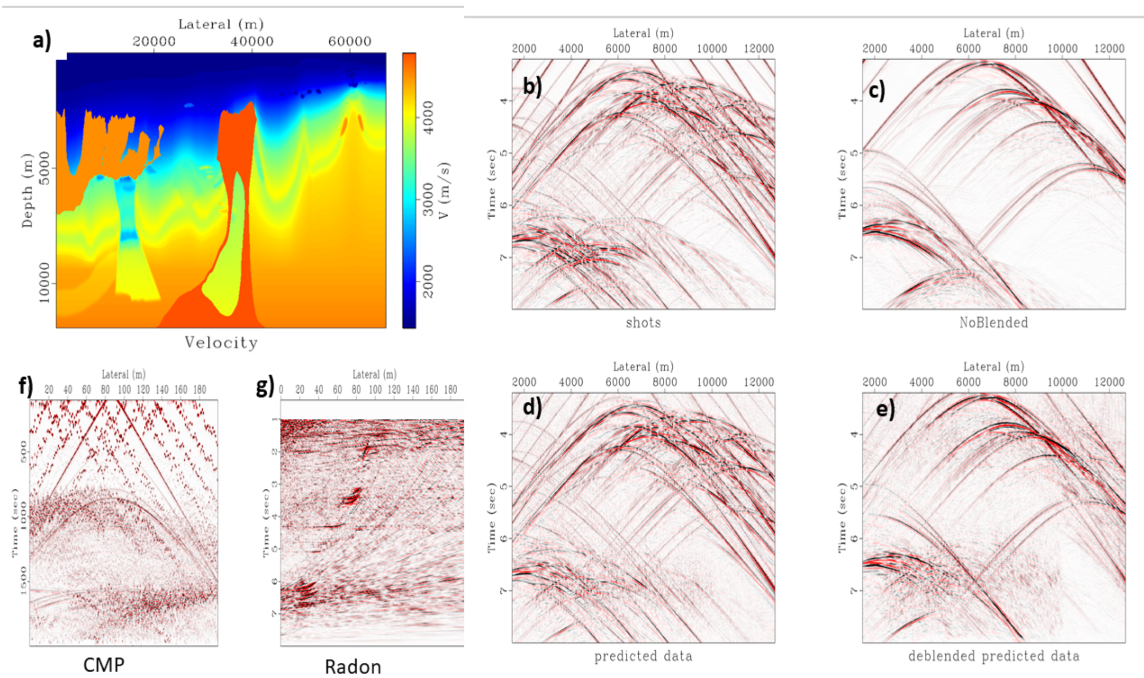

FIG. 9. Deblending by inversion in complex salt structure recording

In all the blending examples so far, we followed a particular blending strategy where shots are separated in distance, which is common, both in marine and land acquisitions. Another approach that is common and perhaps easier to implement in the field is continuous recording. For example, shots can be triggered at short intervals before the energy from the previous shots has dissipated. We will show only a preliminary example since we are still investigating the best way to deblend this case. In Figure 10a a never blended shot modeled in the bp2004 velocity model shown before (Figure 9a). In Figure 10b we see two shots superposed by 50%, Figure 10c shows the result of deblending using inversion in CMPs. Figures Figure 10e and e show one CMP before and after deblending. The separation has some success but some energy appears leaking at the bottom.

In all the blended data sets shown above, there are time delays that we used to turn blended energy into incoherent patterns, either in the CMP or receiver domains. However, we don't always have these delays. The first real data example in Figure 1 did not have delays. In that case, two vibroseis, one on each side of the cable, were triggered simultaneously. This is non-uncommon, for example in a marine acquisition where a boat sailing at the end of the streamer would produce a secondary shot with opposite moveout to the shot at the front (see Trad et al. (2012) for a streamer example in the Gulf of Mexico). How do we address that case?

The main purpose of the delays is to provide the transform with a differentiator, coherent vs incoherent energy. However, this is not the only feature we use in seismic to separate events. One of many other possible differentiators is curvature, as we use in many other problems (for example, multiples). In this case, the seismic events are fitted by a hyperbola, whose position is defined by an apex. Therefore, a reasonable recipe for deblending is to apply different transformations for different apexes and apply sparseness. Figure 11a

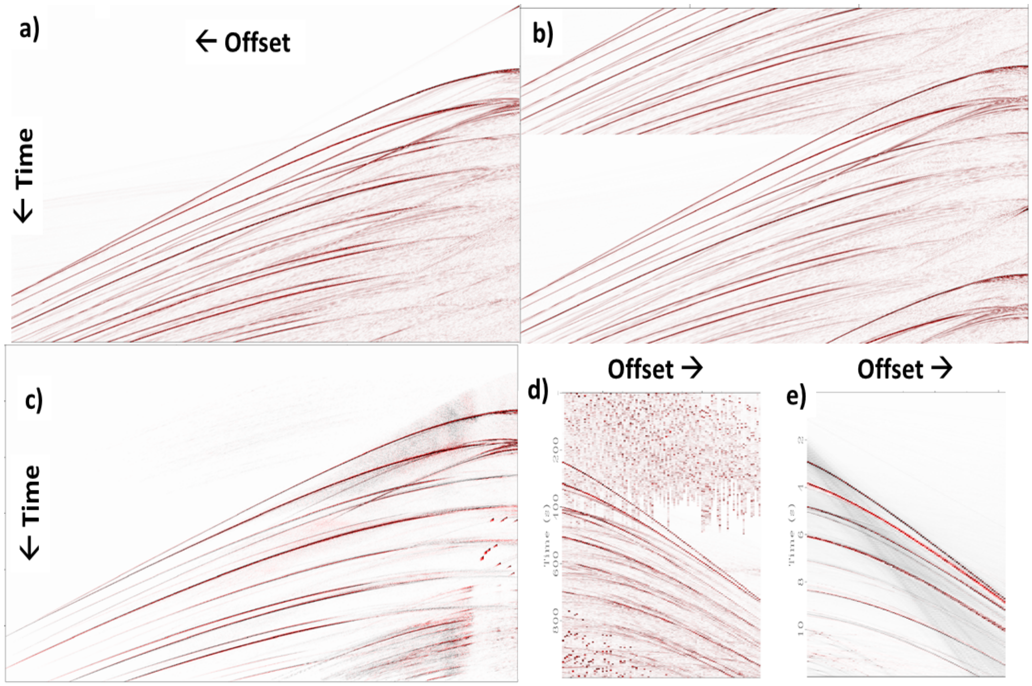

FIG. 10. Deblending by inversion in continuous recording in complex salt structure.

illustrates a hybrid Radon transform (Trad et al., 2001). The idea is that two different transforms fit the data simultaneously. In the most generic case, it is designed to carry two or more operators with different basis, like linear and parabolic. For the case of the CREWES data set, we can use the same basis functions but with different apexes. In Figure 11a we see an example of separating two parabolic events corresponding to two different offsets. Since the offsets are part of the operator (**L**), we need two different operators (Figure 11b). The events are simply separated by applying sparseness on the Radon space Figure 11c, which acts by removing the cross-talk.

Since this deblending technique does not require time delays, we can use if for the CREWES data in Figure 1. In Figures 12a, b and c, we see how the hybrid operator can map both blended shots simultaneously (c) and reproduce them (b) if no action is taken in the Radon space c). Figures 11d, e, f show the separation when the right part of the Radon space f) is muted, predicting only the right dipping events on e). Similarly, Figures 11g, h, i show the prediction of the left dipping events when the left part of the Radon space is mute i).

Deblending and compressive sensing

In the section above, we have seen how by exploiting incoherence and sparseness, we can separate blended shots in different scenarios. In all these cases we chose a hyperbolic Radon transform as one of many possible transforms. In the literature we can see many examples using Fourier transform (Abma et al., 2015), Linear Radon transform (Akerberg et al., 2008), Fx filters (Mahdad et al., 2011), curvelets (Kumar et al., 2015), and others. The name "compressive sensing" (Mosher et al., 2014) has been used often as an umbrella to capture all of these transforms. This name refers to the excellent wide mathematical

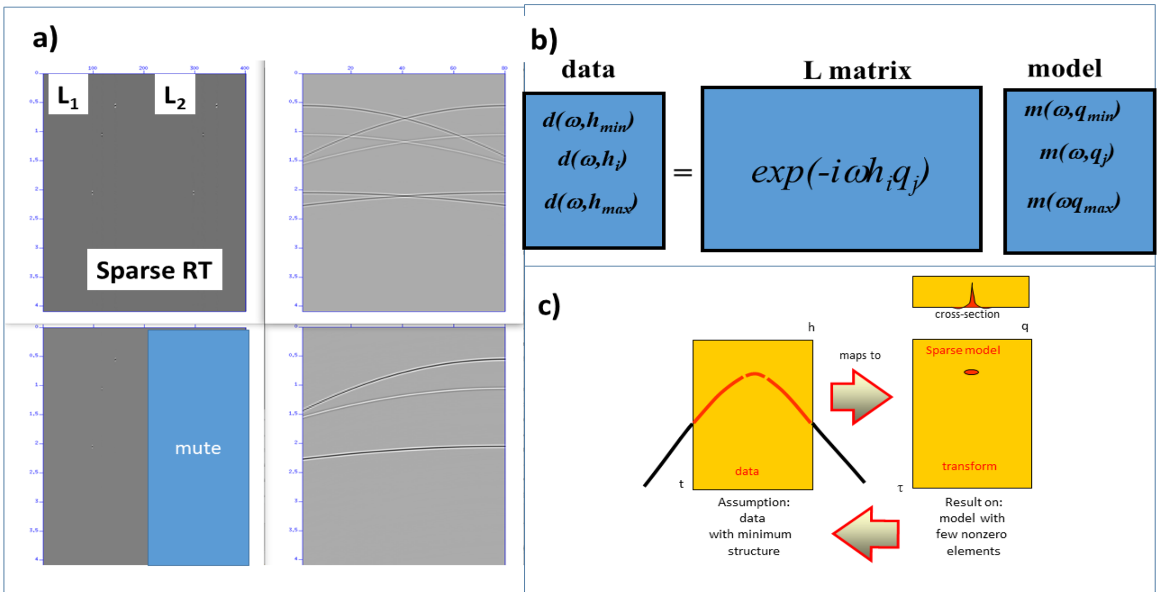

FIG. 11. a) Separation of shots without time dithering. Left, synthetic example where two different operators are used to separate the events. Upper right) Schematics of the Radon frequency domain operator illustrating its dependence on the offset. Bottom right) sparseness principle

framework created by mathematicians in the area of digital sampling (see for example Candès and Wakin (2008)), but these methods can be traced to much older applications in Geophysics by the name of sparse transforms (Claerbout and Muir, 1973; Harlan et al., 1984), and were applied for decades in many geophysical tools like denoising, interpolation and regularization (see Trad (2018) for a mathematical discussion of this equivalence).

Deblending by physical transforms

We have used Hyperbolic Radon transform in our operators, but this choice was one of many possibilities. When applying these transformations normally we try to capture data details in as few elements of information as possible. These elements are coefficients of a transform, and minimizing their number is known as "sparseness" and "maximum entropy". The basis functions of the transforms we have used are in general data independent. We generate them with parametric equations and hope they will capture by simple cross-correlations the data details, so we can separate coherent events from incoherent ones.

This was a common signal processing approach. A question immediately comes to mind: why don't we do better by adapting the basis functions to our data? This is done in some transforms, like singular value decomposition (SVD), and many machine learning techniques. Another way to do this is by using physics. What is the best possible basis function we could have for a particular data set? The answer, perhaps not obvious, is the whole data. Unfortunately, that wouldn't be useful because the data includes energy that we want to separate. We need to use basis functions that we can identify as the construction blocks of the data.

Of the infinite number of possible basis functions that can reproduce the data with very

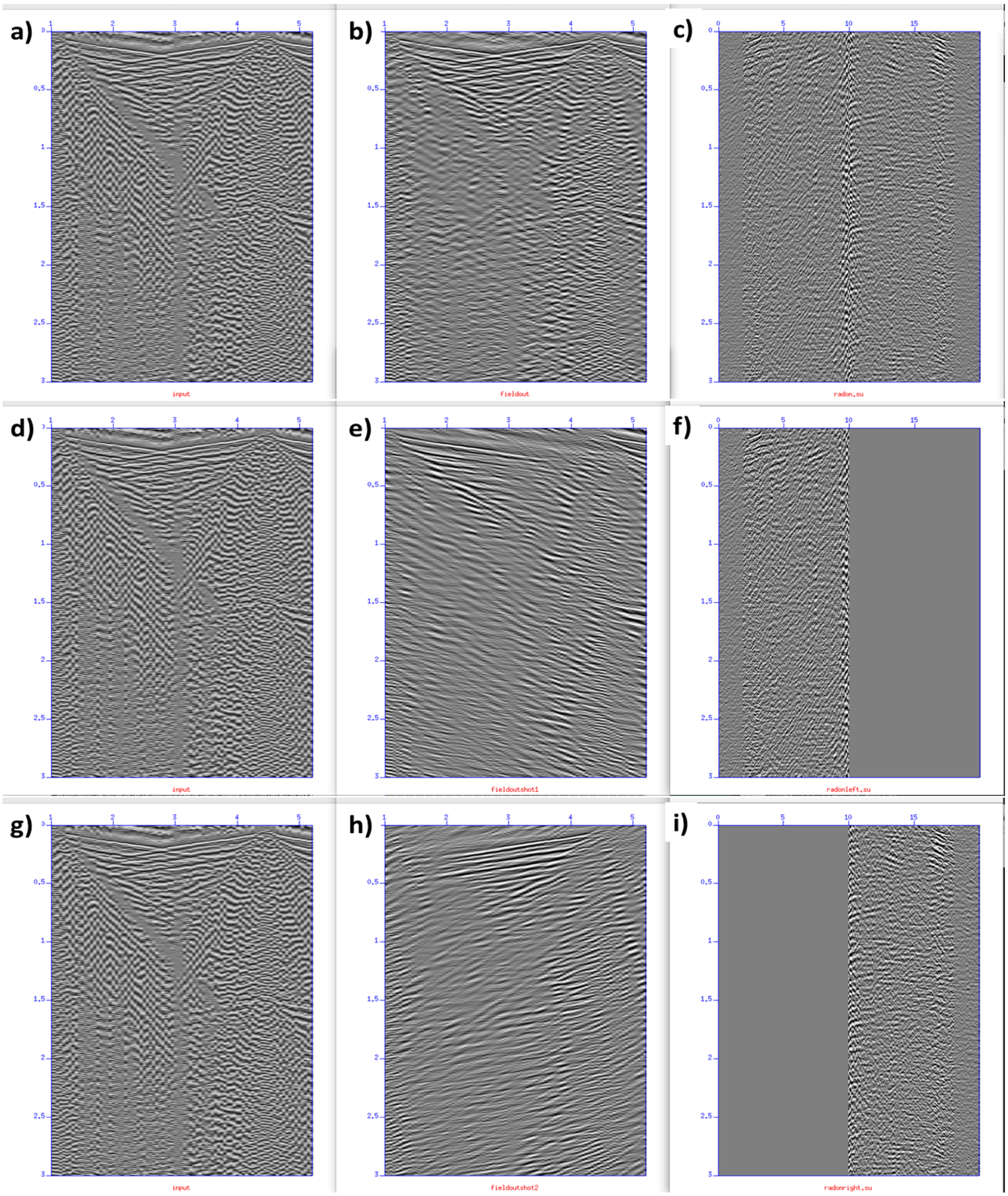

FIG. 12. De-blending of shots acquired by CREWES, (top row) both shots predicted by hybrid Radon transform, (middle row) first shot obtained by muting of right panel, (bottom row) second shot obtained by muting of first panel

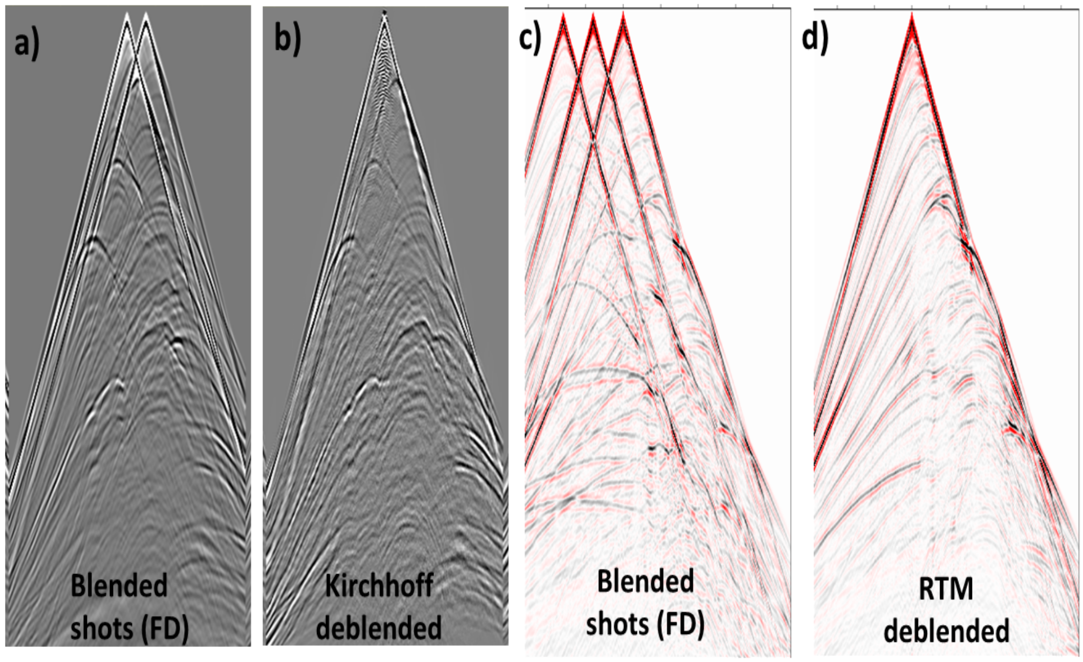

FIG. 13. a) A blended shot produced by finite differences, b) One of the shots predicted by Kirchhoff (migration/demigration) c) Another supershot, this time with three shots, d) deblending by RTM migration/demigration

few coefficients, Green functions are of special interest because they can be obtained with everyday tools: migration/demigration operators. Any migration algorithm is a transform that collapses seismic events to scatter points. The demigration corresponding operator is the inverse transform or synthesis operator that reproduces data from a scatterer point.

Let us see then, several possibilities based on migration/demigration, without forgetting that they are just another type of transformation. Figure 13a shows a two-shot blended gather generated by finite differences using the Marmousi model. Figure 13b shows one of the shots after separation by a migration/demigration Kirchhoff algorithm. In Figure 13c we see three shots from the same model, deblended by prediction with migration/demigration using RTM (Figure 13d). These high-quality results would not be possible in real data with inaccurate velocity models, but they illustrate how migration algorithms can perform very well as transformations for deblending. We call them physical transforms because their basis functions are obtained from physics (wave equation), not by parametric equations.

Seismic is essentially a divide and conquer approach: obtaining an accurate image of the earth for blended data requires that we solve several problems. First, we need to use signal processing to attenuate the noise and remove statics. Secondly, we need to extract velocity information from the blended data through some form of blended-velocity analysis. Third, if we succeed in the previous two, we can think about migration. However, seismic processing is also an iterative approach: we solve partially these three problems and repeat, using the information on each stage to solve the same problems again, every time with more accuracy. The question then is, can we use migration/demigration as a deblending tool after just a partial solution to the previous problems? If we do, then migration/demigration can be added to our package of tools to deblend data.

Usually, we think of migration as a post-processing approach because it requires an accurate velocity model since errors in the velocity model translate to large imaging errors. Data produced by migration/demigration however, does not have the same sensitivity. Kinematic errors in reflectivity created during the migration are compensated by opposite errors in the modelling direction. Dynamic (amplitude) errors, on the other hand, are more persistent, unless we use a data fitting approach like LSMIG. The problem is that LSMIG can be computationally expensive, but many approximations to LSMIG exist that are relatively fast, for example, non-iterative migration.

Looking back at Figure 13 we see two end-members of the LSMIG family: Kirchhoff and RTM. The deblended shot in Figure 13b is less accurate than the shot in Figure 13d but is more robust to velocity errors. These two deblending examples use only one iteration of an LSMIG algorithm. Better accuracy in amplitudes could be achieved by using more iterations. Also, in both cases, we use smooth velocity models. For Kirchhoff, we use ray tracing, but simpler structures could be deblended by time migration directly. The shot in Figure 13d, on the other hand, can not be predicted without a fairly accurate velocity. Both results would degrade with inaccurate velocities but finite-difference methods are more sensitive to velocity errors than ray tracing.

These examples, where the model is a reflectivity model in depth, represent perhaps the best physical transforms, also the most computationally expensive. The advantage of such transformations is that we can use a significant component of our prior geological and geophysical information, that is velocities, densities and reflector positions. However, transforms based on migration/demigration operators can be more useful if we make them less sensitive to the prior information. Figure 14 shows a deblending sequence using the Apex Shifted Radon Transform (Trad et al., 2012; Ibrahim and Sacchi, 2013). The Apex Shifted Radon transform is a modified Least squares Stolt mapping, akin to LSMIG, to collapse hyperbolic events to their apexes (Trad, 2003). This transformation is faster than standard Radon transform, but it works well only on relatively simple scenarios. It is an example of a robust physical transform, but perhaps the low/accuracy member of the family. A more accurate but related operator is the Common Focal transform (CFP) (Berkhout and Verschuur, 2006). It is more expensive but more flexible because it allows one to separate the down and up parts of the wavefield, with interesting opportunities for example for multicomponent processing.

Blended migration

As mentioned by Berkhout (2008) and others, in practice it is possible to get a good image of blended data directly without deblending. Similarly, it is possible to perform Full Waveform Inversion (FWI) without deblending. Regular shots are often blended numerically to migrate plane waves and diminish computational cost. This is a step into the difficult but desirable goal of reducing the number of "divide and conquer" modules that seismic processing requires. Figures 15a-c show a never-blended shot for Marmousi and the RTM image obtained for many of these shots directly. Figures 15b-d show one supershot with five blended shots, and the image obtained by direct migration of the blended

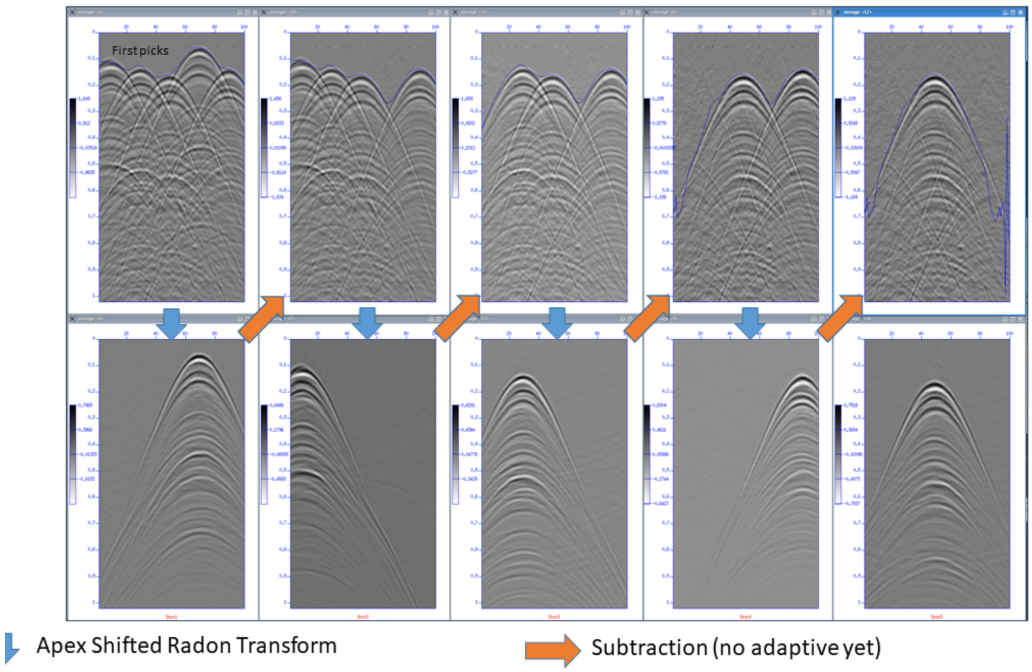

FIG. 14. A five-shot gather deblended by an Apex Shifted Radon Transform sequence

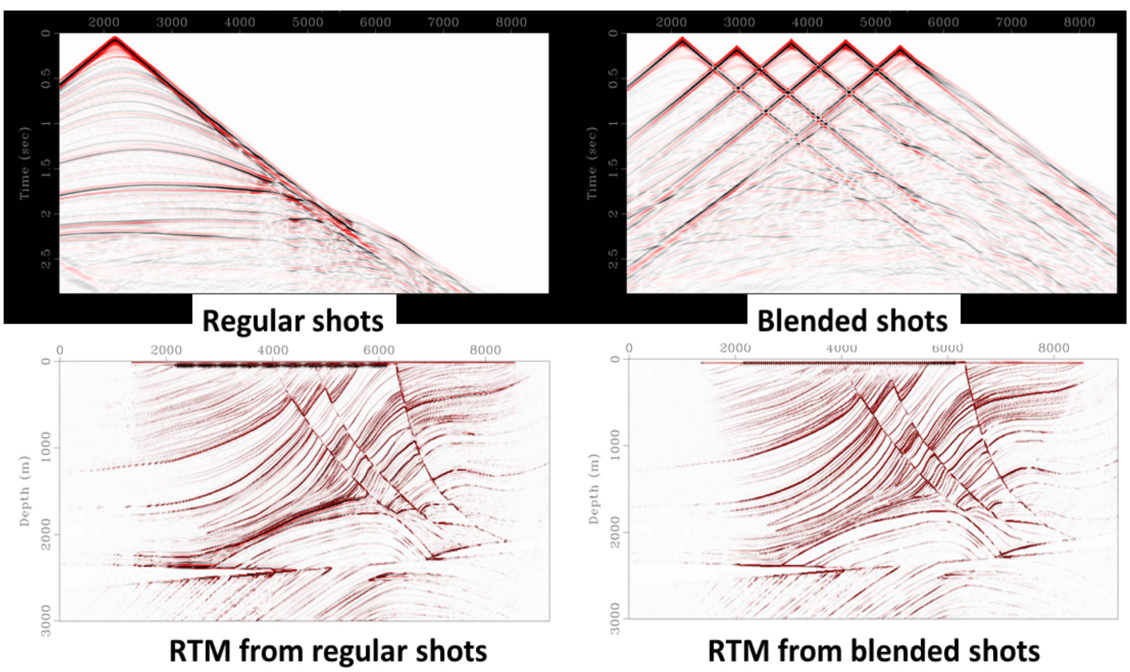

FIG. 15. First column, regular data migrated by RTM. Second column) Five-shot blended data migrated by LSRTM.

data. Of course, this result would not be possible without an accurate velocity model, which still we don't know how to do directly from blended data (FWI would require to start from a very good tomographic model). Therefore, deblending is still needed even if migration can be done without. However, as explained earlier, we can reduce the deblending process as a tool to extract information to use in blended-migration.

SUMMARY AND CONCLUSIONS

In this report, we have seen different approaches to deblending based on current processing techniques, and we have classified them in three types:

- Noise attenuation to eliminate interference: data are replicated as many times as sources are blended, and data sorted in such a way that interference appears incoherent or with different velocities or slopes. Many standard tools like FK and FX filters, Radon transforms, curvelet transforms, and others can be used to remove the interference. The information coming from the blended shots are treated as noise and removed as such. These techniques usually rely on random dithering or time delay introduced between simultaneous shots during acquisition, but not always.
- Multidimensional inversion: the blending process is taken as a mixing matrix or mixing operator and some sort of least-squares inversion with regularization constraints is used to invert this operator to produce a deblended data set. This process can be combined with interpolation/regularization operators. Similarly to the previous technique, this method exploits some time or space perturbation on the shooting which makes the interference to appear incoherent on some spatial direction. At difference of the denoising, in this case, all energy is accounted for.
- Physical transformation: the blended data is transformed into a model with physical (rather than mathematical) meaning. Physical constraints are applied to this model, for example, reflector positions and velocities. The constraints serve as a deblending technique. Data are then predicted from this model into separated shots. Time delays or other means of differentiating shots can be used as well.

As we move from purely mathematical transformations, for example, Fourier as used in 5D interpolation, to physical transforms like migration/demigration and LSMIG, there are more opportunities to use prior information, like velocities. At the same time, prior information errors have negative effects. In between these two ends, purely mathematical transforms and physical transforms, there are intermediate alternatives and combinations of them.

ACKNOWLEDGMENTS

We thank CREWES sponsors for contributing to this seismic research. I also gratefully acknowledge support from NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 461179-13 and NSERC Discovery Grant.

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.

Abma, R., and Yan, J., 2009, Separating simultaneous sources by inversion, *in* 71st EAGE Conference and Exhibition incorporating SPE EUROPEC 2009.

Akerberg, P., Hampson, G., Rickett, J., Martin, H., and Cole, J., 2008, Simultaneous source separation by sparse radon transform, *in* SEG Technical Program Expanded Abstracts 2008, Society of Exploration Geophysicists, 2801–2805.

Beasley, C. J., 2008, A new look at marine simultaneous sources: The Leading Edge, **27**, No. 7, 914–917.

Beasley, C. J., Chambers, R. E., and Jiang, Z., 1998, A new look at simultaneous sources, *in* SEG Technical Program Expanded Abstracts 1998, Society of Exploration Geophysicists, 133–135.

Berkhout, A., and Verschuur, D., 2006, Focal transformation, an imaging concept for signal restoration and noise removal: Geophysics, **71**, No. 6, A55–A59.

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

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

Candès, E. J., and Wakin, M. B., 2008, An introduction to compressive sampling: IEEE signal processing magazine, **25**, No. 2, 21–30.

Claerbout, J., 1992, Earth sounding analysis, Processing versus inversion: Blackwell Scientific Publications, Inc.

Claerbout, J. F., and Muir, F., 1973, Robust modeling with erratic data: Geophysics, **38**, No. 5, 826–844.

Harlan, W. S., Claerbout, J. F., and Rocca, F., 1984, Signal/noise separation and velocity estimation: Geophysics, **49**, 1869–1880.

Ibrahim, A., and Sacchi, M. D., 2013, 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.

Kumar, R., Wason, H., and Herrmann, F. J., 2015, Source separation for simultaneous towed-streamer marine acquisitionâĂŤa compressed sensing approach: Geophysics, **80**, No. 6, WD73–WD88.

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.

Mosher, C., Li, C., Morley, L., Ji, Y., Janiszewski, F., Olson, R., and Brewer, J., 2014, Increasing the efficiency of seismic data acquisition via compressive sensing: The Leading Edge, **33**, No. 4, 386–391.

Trad, D., 2003, Interpolation and multiple attenuation with migration operators: Geophysics, **68**, No. 6, 2043–2054.

Trad, D., 2018, Compressive sensing, sparse transforms and deblending.

Trad, D., Sacchi, M. D., and Ulrych, T. J., 2001, A hybrid linear-hyperbolic radon transform: Journal of Seismic Exploration, **9**, No. 4, 303–318.

Trad, D., and Siliqi, R., 2015, Device and method for de-blending simultaneous shooting data with apex shifted radon transform, uS Patent 9,091,788.

Trad, D., Siliqi, R., Poole, G., Boelle, J.-L. et al., 2012, Fast and robust deblending using apex shifted radon transform, *in* 2012 SEG Annual Meeting, Society of Exploration Geophysicists.

Trad, D. O., Ulrych, T. J., and Sacchi, M. D., 2002, Accurate interpolation with high-resolution time-variant radon transforms: Geophysics, **67**, No. 2, 644–656.

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

Verschuur, E., 2011, Deblending workshop, the hague.