A comparison of shot encoding schemes applied to the Full Waveform Inversion problem

Alan A. V. B. Souza†(*), Andre Bulcão†, Bruno P. Dias†, Djalma† M. S. Filho, Felipe P. Loureiro‡, Fernanda F. Farias†, Gustavo C. Alves†, Luiz A. Santos†, Ricardo F. C. C. da Cruz†, †PETROBRAS S/A, Brazil, ‡Fundação Gorceix

Abstract

This paper compares different shot encoding strategies applied to the FWI (Full Waveform Inversion) problem. The results indicate that these encoding strategies can lead to a decrease of computational cost and or an overall better quality of the final inversion result when compared against FWI without source encoding.

Introduction

One of the main goals of the seismic processing routine is finding a good velocity model that can be used to perform seismic depth imaging or to perform stratigraphic interpretation. Attaining this goal on regions of complex geological settings like the sub-salt of GOM or on Brazilian pre-salt is challenging, because the more conventional techniques or the less computationaly demanding ones can not cope with the complex interactions of the seismic wave-field with the subsurface geological structures.

When the framework of FWI is used, the problem of finding a velocity model is posed as a local optimization problem for a nonlinear objective function that accounts for the changes of amplitude and phase of the field data. This optimization problem is usually solved using a local optimization schemes based on the adjoint method to calculate the gradient and the Hessian of the objective function. For a more in-depth discussion see Tarantola (1988) and Tromp et al. (2005).

Shot encoding schemes appear in the FWI parlance as tools to reduce the humongous computational cost (one iteration of this method is roughly equivalent to one full RTM migration) and to provide in some cases a better signal to noise ratio as a tool to design blended acquisitions.

The encoding schemes presented here are most atractive to FWI methods performed on the time domain, the frequency domain using iterative solvers and using a hybrid approach (where the modeling is done on time domain and inversion on frequency domain, using for example the approach of Nihei and Li (2007)), because in these domains the computational cost is proporcional to the number of gathers necessary to obtain the FWI gradient.

Full waveform inversion problem

Because of the non-linearity of the objective functional the FWI problem is strongly ill-posed,(Virieux and Operto (2009)). To overcome these difficulties, it's common practice to perform the inversion on a multiscale manner on frequency domain and use some kind of regularization.

In this work the inversion was performed on the frequency domain using a modified fourth order finite difference scheme. The multiscale approach and a family of multiplicative regularizators as proposed by Ramírez and Lewis (2010) were utilized. Equation 1, represents the objective functional used:

$$\hat{\mathrm{J}} = \mathrm{J}(\mathbf{u},\mathbf{m})\,\mathrm{R}(\mathbf{m}), \tag{1}$$

where 1 the term $\mathrm{J}(\mathbf{u},\mathbf{m})$ represents a data domain misfit function using L_2 norm and the R term represent the multiplicative regularizer that act on model space. The regularizer employed in this works corresponds to the weighted L_2 regularizer of the Ramirez paper. The model update was governed by the expression 2:

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \alpha \mathbf{s}_k, \tag{2}$$

where s_k and α are suitable search directions and scale parameter, found respectively using a steepest descent techinique and an approximated line search algorithm.

Shot encoding schemes can be applied to the FWI problem (these are also known as mutlisource FWI) because the gradient when are found using the adjoint method resembles a standard RTM migration. More details and application examples can be found on the following articles Ben-Hadj-Ali et al. (2011), Schuster et al. (2011, 2010), Boonyasiriwat and Schuster (2010), Krebs et al. (2009), Vigh and Starr (2008).

Encoding schemes

Encoding schemes are devised as methods for combining shot records in an advantageous form that minimizes the computational cost for imaging this new (encoded) and smaller (in number of gathers) data set and at the same time reduces the cross-talk term that impairs the final image quality (Romero et al. (2000)). In a more general form, encoding schemes can be created manipulating the so called encoding matrix (Godwin (2011); Godwin and Sava (2010)). But in this work the encoding schemes used have the simple form of equation 3:

$$\Phi_j(\omega) = \sum_{i=1}^{n_s} \phi_{ij} S_i = \sum_{i=1}^{n_s} \left[\mathrm{A}_{ij} e^{i\omega\tau_{ij}} \right] S_i, \qquad (3)$$

where the function $\Phi_j(\omega)$ corresponds to the j^{th} *super-gather* of the chosen encoding scheme. Here the encoded gathers are constructed from the original shots by a process of weighted stacking of the original shots controlled by the function ϕ_{ij}. This function has two terms, one of them corresponds to an amplitude weight and the other a time-shift of the shot records. The indexes i and j corresponds to the original i^{th} shot record and j^{th} *super-gather* of the chosen encoding scheme, respectively.

It's usefull to present a brief overview and proper citations for the shot encoding schemes used on this work.

- **Decimated**: Sampled every n^{th} shot from the original survey. Roughly uses $\lfloor \frac{n_s}{n} \rfloor$ gathers instead of the total n_s.
- **Random encoding**: Uses all gathers from the original survey combining all of them after applying to each one a random time delay and an amplitude weighting. The different *super-gathers* are obtained with different realizations of the random time delay and amplitude weigthing. In this work the random enconding uses phase changes (time delays) as in the work of Romero et al. (2000) and the amplitude weighting (polarity changes) as sugested by Dai et al. (2012).
- **Linear Delay**: All the shots from the original survey are delayed acoording to a parameter, normally the horizontal slowness ($\mathbf{p}_j$) at the surface and the distance from a specific surface point. The different gathers are obtained varying the $\mathbf{p}$ parameter as presented byGodwin and Sava (2011).
- **Linear dithered delay**: This schemes adds to the previous one a random perturbation of time proportional to the maximum time delay. These time perturbations lead to interesting interpretations since they can be cast as perturbations to the ray parameter $\mathbf{p}$ used to index the gathers of this encoding scheme, each j^{th} gather correspond to the imaging of a group of plane waves statistically distributed around $\mathbf{p}_j$. More detais about the method can be seen in Perrone and Sava (2010).

The expressions for the ϕ_{ij} terms described on equation 3, can be found on table 1.

The schemes listed on table 1 were used on the inversion results of figure set 4 and graph 5. For the linear delay encoding, the sampling rule used was the same as inZhang et al. (2005). For the dithered linear delay two cases were studied. One of them following the previous rule another one using a constant delay increment. Our intention was to estimate the relative importance of a denser ray parameter sampling versus a broad range of ray parameters, even if the latter uses a lower sampling rate for the ray parameter. This encoding scheme was chosen for this study because each "dithered gather" can sample a range of ray parameters although not uniformly Perrone and Sava (2010).

Encoding Scheme	$\mathrm{A}_{i,j}$	$\tau_{i,j}$
Decimated	$\delta_{i\sigma(j)}$	0
Linear Delay	1	$p_j(x_i - x_{i_m})$
Random Delay	$\mathrm{sgn}\left({}^{(j)}\xi_{1_i}\right)$	${}^{(j)}\xi_{2_i}$
Dithered Linear Delay	1	$\left[p_j + {}^{(j)}\xi_{3_i}\right](x_i - x_{i_m})$

Table 1: Sumary of encoding schemes used (ϕ_{ij} functions). Where δ_{ik} is the kroenecker delta and σ a permutation of the original index i for the shot records. ξ_l with $l \in [1,3]$ represents a uniform random variable with suitable limits. The superscript j indicates the j^{th} realization of the random variable and the subscript i, indicates the the i^{th} sampling of this variable. x_{i_m} is a reference point of the survey and p_j the j^{th} ray parameter.

Results for the Marmousi model

In the following section, we present the results obtained to the inversion of the Marmousi model, Bourgeois et al. (1990), (details of the model and seismic data are found on table 2 and figure 1 shows the true model) using the shot encoding schemes of the previous section. The inversion and finite difference modeling was carried out on frequency domain using a multiscale approach with a initial frequency of 7.32Hz and frequency increments of 1.95Hz up to 22Hz.

The inversion is carried out in two passes. In the first pass the model depicted on figure 2 is used as inpu and after all frequencies are inverted the obtained model is used as input for the second pass. The differents datasets for each shot encoding scheme are displayed on figure 3, the noise shown on these figures is added to the data on time domain and its spectrum is bandlimited roughly matching the spectrum of the modeling wavelet. The level of noise on these figures are representative of the noise level of the input data used on the inversion.

Marmousi information			
Model		Data set	
nx	382	dt	1ms
nz	162	number of shots	96
dx	12m	shot spacing	48m
dz	12m	receiver spacing	12m

Table 2: Description of the Marmousi model and correspondent data set

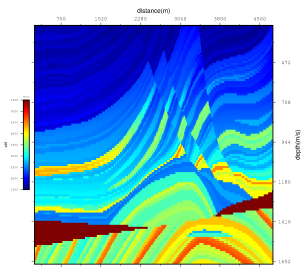


Figure 1: Hard Marmousi model.

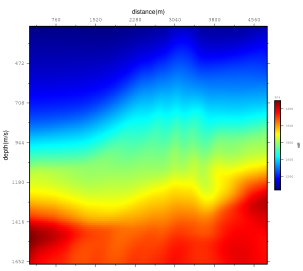
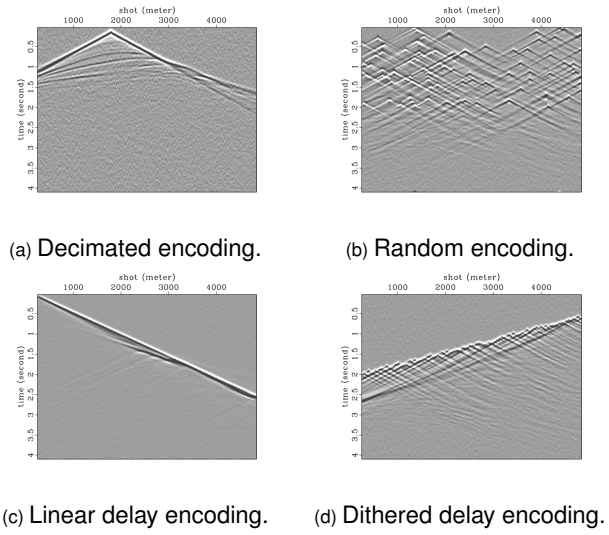


Figure 2: Smoothed Marmousi.

On figure set 4 we have the results for the inversion for different encoding schemes. The leftmost column shows the result with a substantially smaller number of gathers, the center column shows the best result attained by a specific scheme and the rightmost column shows the relative error sections as defined on equation 4 comparing



(a) Decimated encoding. (b) Random encoding.

(c) Linear delay encoding. (d) Dithered delay encoding.

Figure 3: Example of gathers for the different shot encoding schemes used.

the true model with the results of the first column.

$$RE = \left| \frac{m_{true} - m_{inv}}{m_{true}} \right| \quad (4)$$

From figure set 4, first is clearer that in order to recover a good velocity model from the data isn't necessary to use

all available shots. This occurs because of limited model size and because the receivers for all shots are distributed along the horizontal extent of the model. Despite that, it can be seen that the deeper part of the model is less well defined when compared with the upper section. To better reconstruct this part of the model would be necessary to dispose of longer offsets.

The comparison of between shot encoding schemes must be made with caution. "Plane wave" and random encoding schemes use all shot gathers to synthesize its *super-gathers*. This fact can be very usefull, because each of these *super-gathers* illuminate a wider area when compared with the decimated scheme. But the mixing of many gathers can cause an increase of the noise level and compromise the effect of the regularization strategy used.

Discussion

To summaryze the results were created figure 5 and table 3. The former shows important features of the shot encoding schemes such as the rate of SNR increase and maximum attained quality and in the latter displays the parameters used on the constructionn of graph 5.

The abscissa and ordinate displayed in figure 5 are, respectively, the number of *super-gathers* used for each encondig scheme and the signal to noise ratio (SNR) as defined on equation 5:

$$SNR = 10\log_{10}\left(\frac{\sum_z \sum_x m_{true}^2(x,z)}{\sum_z \sum_x [m_{inv}(x,z) - m_{true}(x,z)]^2}\right). \quad (5)$$

This funcion was defined following the paper by Liu et al. (2009) and modified to measure the SNR ratio of an image.

	Decimated	Random	Linear Delay	Dithered Linear Delay	Dithered Linear Delay (constant)
Number of gathers	1-2-4-6-8-12-16-24-48-96	1-2-4-6-8-12-16-24-48-96	1-9-17-25-33-41-49-57-65-73-81-89	1-9-17-25-33-41-49-57-65-73-81-89	1-9-15-19-23-27-31-35-39
Encoding parameter	-	$\tau_{max} = 1s$	$\theta_{max} = 35.41^o$	$\theta_{max} = 35.41^o$	$\theta_{max} = 39.0^o$

Table 3: Table with the shot encoding parameters used. τ_{max} maximum random delay and θ equal to the maximum take-off angle for each set of "plane wave" depicted on the top row, respectively.

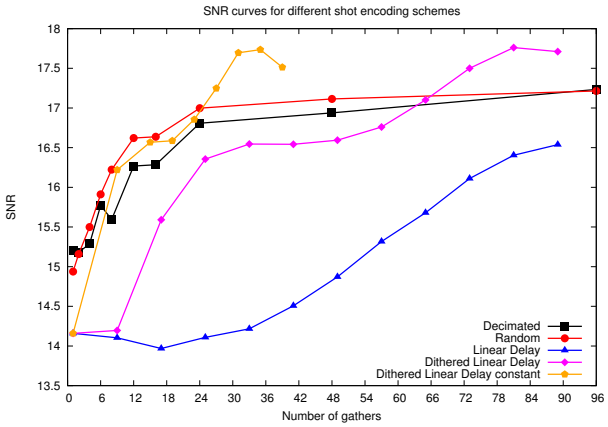


Figure 5: SNR values for different shot encoding schemes and different numbers of gathers used on inversion.

Figure 5 displays the non-linearity of the quality increase for all schemes as function of the numbers of *super-gathers* used. These non-linearities have multiple sources, eg. the type of measurement, a logarithm one. Another source was the resolution limit of the FWI technique. This limitation characterizes the asymptotic behavior of the graph in figure 5. It's possible to understand this behavior observing the localization of the biggest errors on the shallow section of figures 4c, 4f, 4i, 4l, 4o. Those errors are caused by small scales features of the original model like layer boundarires and/or thin layers.

A more subtle source of problems to the model recovery was the regularization strategy adopted in this work. In this strategy the parameter that control the regularization strength (γ on the works of Ramírez and Lewis (2010) and Haffinger (2013)) is dependent on the size of cost function or gradient, as shown in Ramírez and Lewis (2010) and Haffinger (2013). But these functions have magnitudes approximated proportional to the number of *super-gathers* and the quantity of shot records in each one.

The shot encoding schemes used have the property of increasing the numerical value of both gradient norm and cost function, weakening the effect of the regularization. This effect can be seen by comparing the results for the decimated scheme (figure 4b) and plane wave like and random encoding schemes (figures 4e, 4h, 4k and 4n). The result for the decimated scheme looks more blocky, continuous and less noisy when compared to the others schemes. This phenomenon can be explained by the fact that the γ parameter was reduced in the last case due to

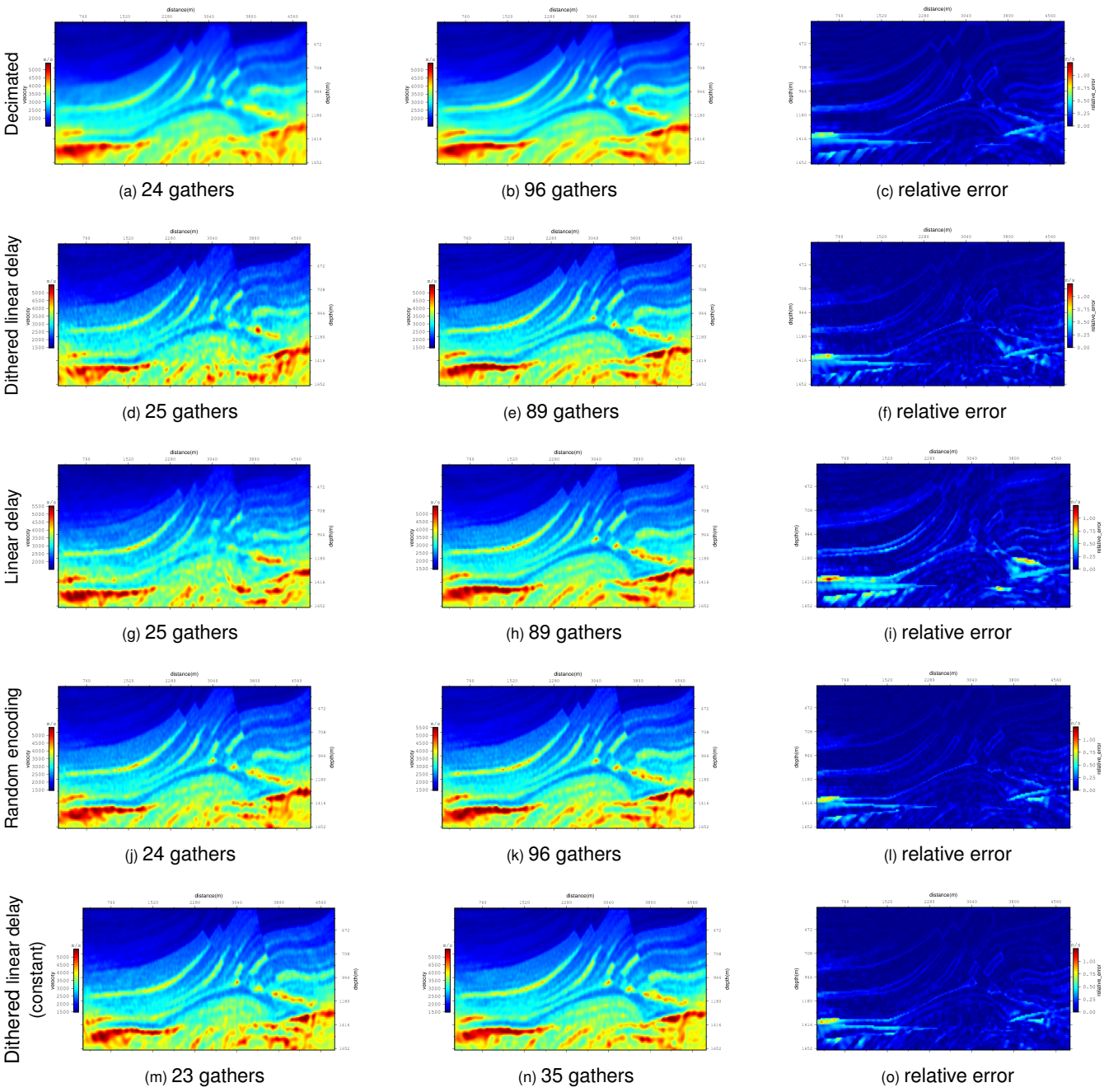


Figure 4: FWI results for various shot encoding schemes. From top to bottom row results for decimated, dithered linear delay, linear delay and random encoding schemes.

the greater magnitude of the cost function.

For the plane wave like encoding schemes the studies realized indicate that the preponderant factor to the quality of inversion was a wide and diversified range of ray parameters. Justifying the difference between the results for the linear delay when comparared against the dithered linear delay. The later effectively use more ray parameters to perform the imaging as discussed on section about the encoding schemes.

The second experiment with the dithered linear delay does not follwo the samplig criteria proposed on Zhang et al. (2005). The experiments shows that a greater range of ray parameters is desirable instead of denser sampling on the $\mathbf{p}$ parameter. This fact can be used to explain the delay of plane wave encoding in improving your own results, because the range of ray parameter that can be imaged using the sampling criteria proprosed in Zhang et al. (2005) grow slowly whith the number of *super-gathers* used.

Finally, the results for the random enconding scheme when compared to the decimated scheme are in agreement with results presented in Schuster et al. (2010, 2011). These articles concludes that the SNR grows with the number of migrations/iterations performed. When applied to the FWI, these results indicate that is necessary to perform a number of "migrations" equal to the number of shot gathers that compose the *super-gather* in order to obtain a simmilar level of signal to noise ratio (Ben-Hadj-Ali et al. (2011)).

Conclusion

The studies show that shot encoding schemes can be used as tool to reduce the computational burden of FWI problems (mainly for time domain, hybrid implementations and frequency domain implementations using iterative solvers). But the choice of the encoding method can be tricky because the inversion results for the same shot encoding scheme can change a lot depending on the parametrization of the choosen encoding. Further investigations regarding the determination of practical rules for the shot encoding scheme are necessary as well as studies of another forms of encoding schemes.

geophysicist

Acknowledgments

We would like to thank PETROBRAS for permission to publish this work.

References

Ben-Hadj-Ali, H., S. Operto, and J. Virieux, 2011, An efficient frequency-domain full waveform inversion method using simultaneous encoded sources: GEOPHYSICS, **76**, R109–R124.

Boonyasiriwat, C., and G. T. Schuster, 2010, 3d multisource full-waveform inversion using dynamic random phase encoding: SEG Technical Program Expanded Abstracts 2010, 1044–1049.

Bourgeois, A., B. M., Lailly, P. Poulet, M. Ricarte, and V. R., 1990, Marmousi, model and data: Presented at the Proceedings of 1990 EAGE workshop on practical aspects of seismic data inversion.

Dai, W., P. Fowler, and G. T. Schuster, 2012, Multi-source least-squares reverse time migration: Geophysical Prospecting, **60**, 681–695.

Godwin, J., and P. Sava, 2010, Simultaneous source imaging by amplitude encoding: Technical Report CWP-645, Colorado School of Mines.

———, 2011, A comparison of shot encoding schemes for wave-equation migration: Technical Report CWP-683, Colorado School of Mines.

Godwin, J. A., 2011, Blended source imaging by amplitude encoding: Master thesis, Colorado School of Mines.

Haffinger, P. R., 2013, Seismic broadband full waveform inversion by shot/receiver refocusing: PhD thesis, TU Delft, Nederland.

Krebs, J., J. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M. Lacasse, 2009, Fast full-wavefield seismic inversion using encoded sources: GEOPHYSICS, **74**, WCC177–WCC188.

Liu, G., S. Fomel, L. Jin, and X. Chen, 2009, Stacking seismic data using local correlation: GEOPHYSICS, **74**, V43–V48.

Nihei, K. T., and X. Li, 2007, Frequency response modelling of seismic waves using finite difference time domain with phase sensitive detection (td—psd): Geophysical Journal International, **169**, 1069–1078.

Perrone, F., and P. Sava, 2010, Wave-equation migration with dithered plane-waves: Technical report, Colorado School of Mines.

Ramírez, A. C., and W. R. Lewis, 2010, Regularization and full-waveform inversion: A two step approach: SEG Technical Program Expanded Abstracts 2010, 2773–2778.

Romero, L., D. Ghiglia, C. Ober, and S. Morton, 2000, Phase encoding of shot records in prestack migration: GEOPHYSICS, **65**, 426–436.

Schuster, G. T., W. Dai, G. Zhan, and C. Boonyasiriwat, 2010, Theory of multisource crosstalk reduction by phase encoded statics: SEG Technical Program Expanded Abstracts 2010, 3110–3114.

Schuster, G. T., X. Wang, Y. Huang, W. Dai, and C. Boonyasiriwat, 2011, Theory of multisource crosstalk reduction by phase encoded statics: Geophysical Journal International, **184**, 1289–1303.

Tarantola, A., 1988, Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation: Pure and Applied Geophysics, **128**, 365–399.

Tromp, J., C. Tape, and Q. Liu, 2005, Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels: Geophysical Journal International, **160**, 195–216.

Vigh, D., and E. Starr, 2008, 3d prestack plane-wave, full-waveform inversion: GEOPHYSICS, **73**, VE135–VE144.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: GEOPHYSICS, **74**, WCC1–WCC26.

Zhang, Y., J. Sun, C. Notfors, S. Gray, L. Chernis, and J. Young, 2005, Delayed-shot 3d depth migration: GEOPHYSICS, **70**, E21–E28.