Groundroll prediction by interferometry and separation by curvelet-domain matched filtering

Jiupeng Yan and Felix J. Herrmann, SLIM, EOS-UBC

SUMMARY

The removal of groundroll in land based seismic data is a critical step for seismic imaging. In this paper, we introduce a work flow to predict the groundroll by interferometry and then separate the groundroll in the curvelet domain. Thus workflow is similar to the workflow of surface-related multiple elimination (SRME). By exploiting the adaptability and sparsity of curvelets, we are able to significantly improve the separation of groundroll in comparison to results yielded by frequency-domain adaptive subtraction methods. We provide synthetic data example to illustrate our claim.

INTRODUCTION

In most land seismic surveys, the seismic sources produce various surface waves depending on the near surface environment. These surface waves, also known as groundroll, can dominate the record and obscure important reflector information. Similar to the removal of surface related multiples, the removal of groundroll consists of two parts: prediction and separation. For the prediction, an interferometry method has been recently proposed by Halliday et al. (2007) and followed by others including Vasconcelos et al. (2008); Dong et al. (2006). For the second separation step, adaptive-subtraction methods, which are used during surface-related multiple elimination (SRME) (Verschuur et al., 1992), have been used successfully by Halliday et al. (2007); Vasconcelos et al. (2008); Dong et al. (2006).

In this abstract, we put the interferometry prediction into a scheme similar to the SRME method. We use the data itself as an operator. Instead of predicting the multiples by multi-dimensional convolution, as in the SRME method, we predict the groundroll by multi-dimensional correlations with a restriction mask. For the separation, Herrmann et al. (2007) has developed the curvelet-domain matching filter and successfully applied it to the primary-multiple separation. In addition, a Bayesian separation algorithm was introduced by Saab et al. (2007); Wang et al. (2008) to improve the overall separation of multiples after matching. This Bayesian separation method was later applied by Yarham and Herrmann (2008) for the purpose of groundroll separation. These authors demonstrate that the Bayesian algorithm can be used to improve the results of signal separation problems where the prediction for the to-be-separated signals is not perfect.

Aside from applying the Bayesian separation method on the groundroll predictions obtained with interferometry, the main point of this paper is to add curvelet-domain matched filtering after applying a conservative Fourier matching for each offset separately.

After this matching, we improve the overall separation by using the Bayesian separation method. The results show that we have achieved good improvements for the groundroll separation problem.

PREDICTION OF GROUNDROLL

Wapenaar and Fokkema (2006) has shown that interferometry of elastic waves could be achieved by cross-correlations:

$$2\mathbb{R}\{G^{v,f}_{p,g}(\mathbf{r_A},\mathbf{r_B})\} \approx \frac{2}{\rho c_K}\int_{\partial V} G^{v,\Phi}_{q,K}(\mathbf{r_A},s)G^{v,\Phi}_{q,K}(\mathbf{r_B},s)^* d\mathbf{s} \quad (1)$$

where G is the Green's function in the frequency domain, the superscript v indicates that the measured quantity is the particle velocity, the subscript p or q indicates the direction of the particle velocity, the superscript denotes the type of the source, such as P-wave source (with $K=0$) or S-wave source(with $K=1,2,3$ indicating different polarization direction), c_K is the velocity of different waves. $\mathbb{R}$ means taking the real part.

Equation 1 shows that by integrating (i.e. summing) the correlation of responses at $r_{A,B}$, from physical sources on surface ∂V, we can produce the full wavefield response $G(\mathbf{r_A},\mathbf{r_B})$ at r_A from a pseudo source at r_B. The key idea of interferometry groundroll prediction (Halliday et al., 2007) is that when summing over sources at ∂V, every physical source is a stationary source for the groundroll, while only a few physical sources are stationary sources for the reflections. Among these few stationary sources, the physical sources at $r_{A,B}$ contribute to the reflections the most. Thus, we can construct the surface wave response $G_{surface}(\mathbf{r_A},\mathbf{r_B})$ at r_A, by muting stationary sources for reflected waves during the integration i.e. by changing the integration interval ∂V in Equation 1 to $\partial V_{surface}$ which is a subset of ∂V containing sources contribute mainly to surface wave responses. A more detailed explanation has been showed by Vasconcelos et al. (2008). Dong et al. (2006) also showed that even if integrating over all the sources according to Equation 1, the surface waves have been relatively amplified compared to the reflections in the full response $G(\mathbf{r_A},\mathbf{r_B})$.

If we put the data into a frequency-receiver-shot data volume $\boldsymbol{P}$ as done within SRME (Verschuur and Berkhout, 1997), the prediction of the groundroll becomes:

$$\breve{\mathbf{g}} = \boldsymbol{RPP}^*, \quad (2)$$

where $\breve{\mathbf{g}}$ is the predicted groundroll, and $\boldsymbol{P}^*$ means conjugate transpose of $\mathbf{P}$ (for each frequency). $\boldsymbol{R}$ is a picking operator. In our case, $\boldsymbol{R}$ is a matrix with zeros on the diagonal and ones elsewhere. Consequently, $\boldsymbol{R}$ picks only sources that contribute to the surface wave response. In conclusion, we use $\boldsymbol{RP}$ as a groundroll prediction operator on the data $\boldsymbol{P}^*$.

SEPARATION OF GROUNDROLL

After we predict the groundroll, the next step is to subtract the prediction from the full wavefield to produce the reflections. In this interferometry prediction process, the prediction is the correlation of two wavefields. The predicted amplitudes and arrival times of groundroll are not perfect. So we have to subtract the groundroll adaptively instead of directly. In other words, we need to match our groundroll prediction to the actual

groundroll in the data before we subtract it. Vasconcelos et al. (2008) showed two ways to adaptively subtract the correlation-groundroll prediction by using the least-squares subtraction technique. One is to subtract the prediction from the original full data, the other is to subtract from the pseudo full-wavefield response produced by Equation 1. According to Dong et al. (2006), groundroll has already been amplified compared to the reflections in pseudo full-wavefield response, i.e. in the pseudo full-wavefield responses, the groundroll become more dominant and we lose reflection information. In order to get a better result, we choose to subtract the prediction from the original full data.

Curvelet-domain matched filtering

Instead of only using the Fourier domain least-square subtraction techniques for each offset, we use an additional curvelet-domain matched filtering technique that was proposed by Herrmann et al. (2007) and has been successfully applied to the removal of surface-related multiples.

According to Herrmann et al. (2007), this curvelet-domain matched filter is based on the following assumptions : (i) the stationary difference is removed by a global (Fourier) matching procedure, which corresponds to removal of the source/receiver directivity during primary-multiple separation (Herrmann et al., 2008b), and (ii) the remaining non-stationary difference is assumed to vary smoothly in phase space—i.e., the amplitude mismatches are assumed to vary smoothly as a function of position and dip along coherent wavefronts.

For the groundroll separation problem, we use the data itself as a groundroll prediction operator which is similar to the SRME. The difference is that the convolution process in SRME produces the multiple prediction while the correlation process in our method amplifies groundroll and suppresses the reflections to produce the groundroll prediction. Thus the above assumptions for matching multiples should stay the same for matching groundroll. The stationary differences in the multiples come mainly from the convolution of the source signature while in groundroll they mainly come from the correlation of source signature. The remaining non-stationary differences in multiples and groundrolls both come from the non-ideal circumstances in the prediction.

To get rid of the stationary difference which is mainly due to the correlation of the source wavelet, a global Fourier matching procedure is applied:

$$\min_{\mathbf{a}} \|\mathbf{a} * \breve{\mathbf{g}} - \mathbf{f}\|^2, \tag{3}$$

where $\mathbf{a}$ is a short (hence smooth in the Fourier domain) global time-domain matching filter, and $\mathbf{f}$ is the original data. $*$ indicates the convolution operation. Considering correlation as just the conjugate operation to convolution, Equation 3 can remove the imprint of the correlated source wavelet. The optimization problem in Equation 3 is solved by the LSQR solver (Paige and Saunders, 1982).

After globally removing the stationary difference , the argument can be made that the non-stationary operator that links the predicted groundroll to the actual groundroll can be modeled mathematically by a zero-order pseudo-differential operator (ΨDO). This operator, some sort of position and dip-dependent filter, can be approximated by a positive scaling in the curvelet domain (Herrmann et al., 2008b):

$$\widetilde{\mathbf{g}} \approx \boldsymbol{C}^T \operatorname{diag}\{\mathbf{w}\}\boldsymbol{C}\mathbf{f}, \quad \{w\}_{\mu\in\mathscr{M}} > 0, \tag{4}$$

where $\widetilde{\mathbf{g}}$ is the matched groundroll, $\mathbf{f}$ is the full wavefield, $\boldsymbol{C}$ is the 2D curvelet transform, $\mathbf{w}$ is the curvelet-domain scaling vector and $\mathscr{M}$ is the index set of curvelet coefficients.

In this approximate forward model, the two wavefields are matched by a simple scaling of the curvelet coefficients. Since the curvelet transform is a higher dimensional generalization of the Wavelet transform designed to represent images at different scales and different angles (Candés et al., 2006), the scaling of the curvelet coefficients gives a position, scale and angle dependent amplitude correction. This is the basis for our curvelet-domain matched filtering.

Based on Equation 4, in order to match our predicted (and globally matched) groundroll to the right scale, we need to estimate a positive curvelet-domain scaling vector $\mathbf{w}$. Since we have the full wavefield and the predicted (and globally matched) groundroll, we estimate the curvelet-domain scaling vector by minimize least-squares mismatch between these two.

The following issues complicate the estimation of the scaling vector: (i) the undeterminedness of the forward model, due to the redundancy of the curvelet transform—i.e., $\boldsymbol{C}\boldsymbol{C}^T$ is rank deficient; (ii) the risk of overfitting the data, and (iii) the positivity requirement for the scaling vector. To address issues (i-ii), we add a smooth constrain to our minimization formulation, which can be expressed as an augmented system of equations:

$$\begin{bmatrix}\mathbf{g}\\ \mathbf{0}\end{bmatrix} = \begin{bmatrix}\boldsymbol{C}^T \operatorname{diag}\{\boldsymbol{C}\breve{\mathbf{f}}\}\\ \gamma\boldsymbol{L}\end{bmatrix}\mathbf{w} \tag{5}$$

i.e. $\mathbf{d} = \boldsymbol{F}_\gamma \mathbf{w}$, where $\mathbf{d}$ is the augmented vector in the left side of Equation 5. The scaling vector is found by minimizing the objective function, a combination of misfit and smoothness in the curvelet domain:

$$J_\gamma(\mathbf{z}) = \frac{1}{2}\|\mathbf{d} - \boldsymbol{F}_\gamma e^{\mathbf{z}}\|_2^2 \tag{6}$$

To guarantee positivity of the scaling vector (issue(iii)), we substitute $\mathbf{w}$ with $e^{\mathbf{z}}$ (Vogel, 2002). Our formulation seeks a solution fitting the vector, $\mathbf{g}$, with a smoothness constraint imposed by the sharpening operator $\boldsymbol{L}$. (Herrmann et al., 2008a). The amount of smoothing is controlled by the parameter γ. For a larger γ, there is more smoothness at the expense of overfitting the data (e.g., erroneously fitting the reflections).

Bayesian separation algorithm

Ideally, the scaled groundroll, yielded by the solution of above nonlinear least-squares problem, $\widetilde{\mathbf{z}} = \arg\min_{\mathbf{z}} J(\mathbf{z})$, could be subtracted from the total data directly. Unfortunately, the noises in seismic data and imperfections in the prediction may affect our curvelet-domain filtering, making the direct subtraction ineffective. The same problem arises in the primary-multiple

separation(Herrmann et al., 2008a). Due to the large amplitudes of groundroll, the imperfection in the matched groundroll could be a more severe problem, i.e. a relatively small amplitude imperfection in the matched groundroll could still cover important reflection information. Recently Saab et al. (2007); Wang et al. (2008) developed a Bayesian separation algorithm to address this problem and successfully applied to multiples. Yarham and Herrmann (2008) applied this Bayesian method to improve the groundroll-reflection separation when a prediction for groundroll from conventional methods is provided.

Following their work, we separate the groundroll and reflections by solving the sparsity-promoting program below:

$$\begin{aligned}\{\widetilde{\mathbf{x}}_1, \widetilde{\mathbf{x}}_2\} &= \arg\min_{\mathbf{x}_1,\mathbf{x}_2} \lambda_1 \|\mathbf{x}_1\|_{1,\mathbf{w}_1} + \lambda_2 \|\mathbf{x}_2\|_{1,\mathbf{w}_2} \\ &+ \|\boldsymbol{C}^T\mathbf{x}_2 - \mathbf{b}_2\|_2^2 + \eta \|\boldsymbol{C}^T(\mathbf{x}_1+\mathbf{x}_2) - \mathbf{b}\|_2^2, \quad (7)\end{aligned}$$

The vectors $\{\widetilde{\mathbf{x}}_1, \widetilde{\mathbf{x}}_2\}$ are the estimates for the reflections and groundroll in the curvelet domain, i.e. the solution we are seeking for. $\boldsymbol{C}^T$ is the inverse curvelet transform. $\{\mathbf{w}_1, \mathbf{w}_2\}$ are positive weights, and $\{\mathbf{b}, \mathbf{b}_2\}$ are the total data and the (curvelet-domain matched) groundroll prediction. In this Bayesian formulation, both the groundroll prediction and the weights depend on the curvelet-domain matching procedure. Our Bayesian formulation can be viewed as an optimization problem. The minimization of Equation 7 is a combination of minimizing weighed l_1 norm of curvelet coefficient, the $l_2 - misfit$ of the matched and predicted groundroll, and the $l_2 - misfit$ of the sum of estimated groundroll and reflections and the total observed data. Curvelet transform and l_1 minimization utilize and promote the sparsity of seismic data, i.e. the sparsity constrain of $\mathbf{x}_1, \mathbf{x}_2$ increases the chance of separating and makes the solution more 'seismic'. The combined $l_2 - misfit$ allows us to choose a confidence level of our prediction. By selecting the parameter λ's and η , we control the weighting of different components of our objective function.

NUMERICAL EXAMPLE

In this section, we provide a synthetic example for our workflow. The data is produced by using the velocity model in Fig 1. A line of receivers is horizontally placed on the surface with 8m space distance. Both vertical and horizontal point sources are placed at the same position of receivers one after another to produce the shot gathers. Although we tested our methods on all components and source types, here we only show the results from the vertical components of the vertical sources, which groundroll is most dominant.

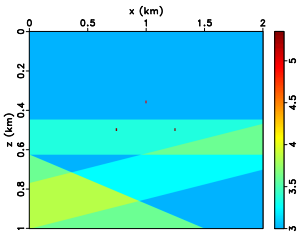

Figure 1: P-wave velocity model.

Fig 3(a) shows a synthetic shot gather at 1.0km. Fig 3(b) shows the true reflections of a shot at 1.0km. Fig 3(c) shows the correlation interferometry prediction of groundroll. Fig 3(d) shows the reflections produced from our separation workflow. Fig 2 shows the results for the shot gather at 0.4km. Fig 4 shows the results for the shot gather at 1.6km.

From current results, we can see that our separation method successfully removes most of the groundroll. In the original full data, reflections are obscured by strong groundroll that can barely be seen. Our separation results capture nearly all reflection information, although there is still some groundroll residual left. We think this is due to overfitting of the data during the matching. There are also numerical artifacts in the modeling. Even though we achieved reasonably good results, a more thorough examination of the parameters in the global Fourier matching and curvelet-domain matching process may yield even better results. Further more, we intend to apply our method to real data.

DISCUSSION

In this paper, we present a comprehensive workflow to predict and separate the groundroll in seismic land data. Our workflow is similar to the one successfully employed during surface-related multiple elimination. Table 1 compares our workflow with SRME. These coherent-noise removal methods consist of two steps, namely, a prediction and a separation step. Each of these steps is important and the performance of both steps are interrelated, e.g., the separation by the Bayesian algorithm will be unsuccessful when the predictions are over fitted. Even in cases where the predictions are reasonable, e.g., during multiple removal, Herrmann et al. (2008b) showed that significant improvements can be made by exploiting curvelet-domain adaptation and sparsity. Our findings in this paper are similar, event though the removal of groundroll is more difficult because of the large amplitudes and dispersive behavior of groundroll. However, our results on synthetic data show that the curvelet-based method of matching and Bayesian separation can lead a better removal of groundroll and a better preservation of the reflections, .

	SRME for multiples	our method for groundroll
prediction	$\boldsymbol{P}_0\boldsymbol{P}$	$\boldsymbol{RPP}^*$
separation	global Fourier matching local Fourier matching	global Fourier matching curvelet matched filtering
improving	iterative SRME	Bayesian separation

Table 1: Comparison of our groundroll separation method with SRME method. $\boldsymbol{P}_0$ is the primary operator Berkhout and Verschuur (1997) which is not available in first iteration of SRME and substituted by total data operator $\boldsymbol{P}$.

ACKNOWLEDGMENTS

The authors would like to thank Ivan Vasconcelos and ION-GXT for providing us the synthetic data. This work was in part financially supported by the NSERC Discovery Grant (22R81254) of F.J.H. and CRD Grant DNOISE (334810-05).

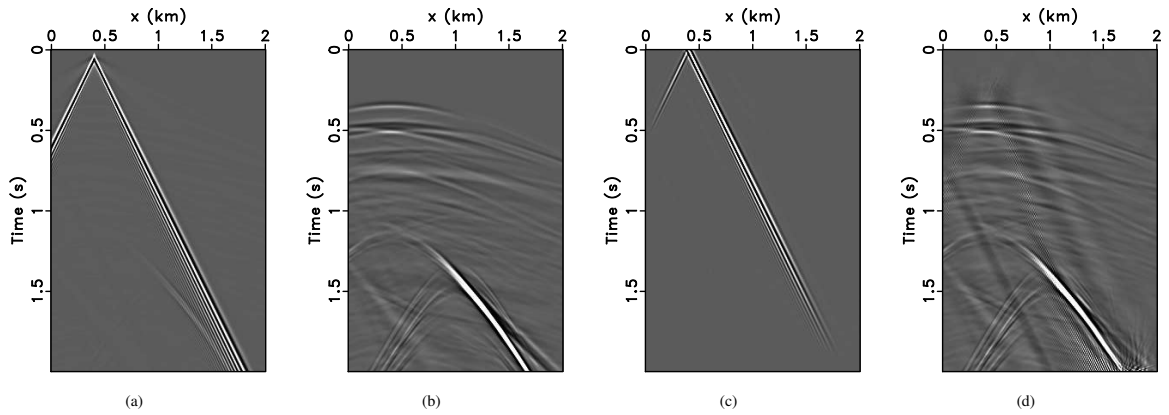

Figure 2: (a) Vertical component of particle-velocity of a vertical source at 0.4km. (b) True reflections contained in (a). (c) Interferometry prediction of the groundroll in (a). (d) Reflections produced from our separation workflow. (d) is produced by adaptively subtracting (c) form (a).

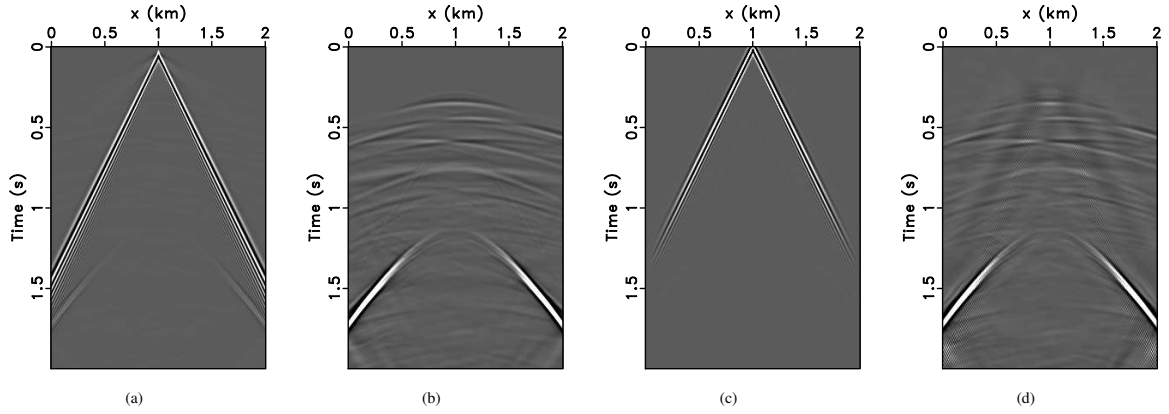

Figure 3: (a) Vertical component of particle-velocity of a vertical source at 1.0km. (b) True reflections contained in (a). (c) Interferometry prediction of the groundroll in (a). (d) Reflections produced from our separation workflow. (d) is produced by adaptively subtracting (c) form (a).

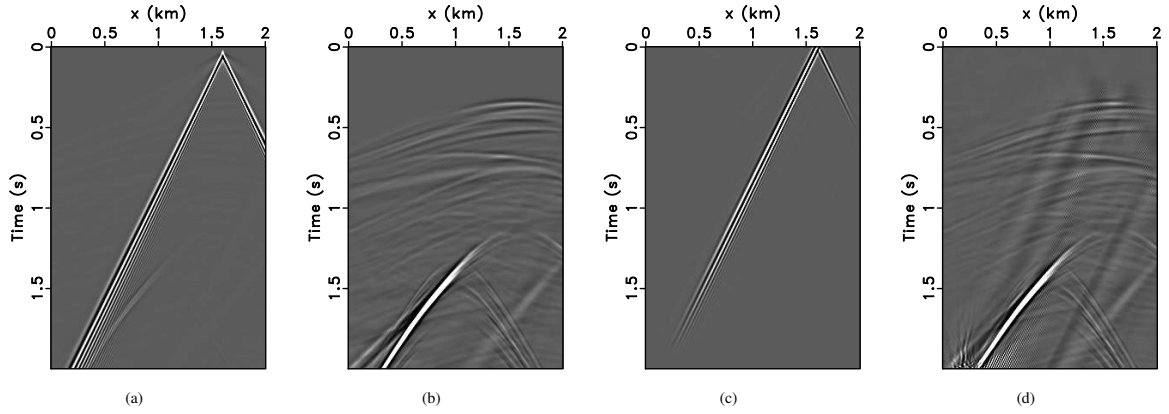

Figure 4: (a) Vertical component of particle-velocity of a vertical source at 1.6km. (b) True reflections contained in (a). (c) Interferometry prediction of the groundroll in (a). (d) Reflections produced from our separation workflow. (d) is produced by adaptively subtracting (c) form (a).

REFERENCES

Berkhout, A. J., and D. J. Verschuur, 1997, Estimation of multiple scattering by iterative inversion, part i: Theoretical considerations: Geophysics, **62**, 1586–1595.

Candés, E. J., L. Demanet, D. L. Donoho, and L. Ying, 2006, Fast discrete curvelet transforms: SIAM Multiscale Model. Simul., **5**, 861–899.

Dong, S., R. He, and G. T. Schuster, 2006, Interferometric predcition and least squares subtraction of surface waves: SEG Technical Program Expanded Abstracts, 2783–2786, SEG.

Halliday, D. F., A. Curtis, J. O. A. Robertsson, and D.-J. van Manen, 2007, Interferometric surface-wave isolation and removal: Geophysics, **72**, A69–A73.

Herrmann, F. J., U. Boeniger, and D. J. Verschuur, 2007, Non-linear primary-multiple separation with directional curvelet frames: Geophysical Journal International, **170**, no. 2, 781–799.

Herrmann, F. J., P. P. Moghaddam, and C. C. Stolk, 2008a, Sparsity- and continuity-promoting seismic imaging with curvelet frames: Journal of Applied and Computational Harmonic Analysis, **24**, 150–173. (doi:10.1016/j.acha.2007.06.007).

Herrmann, F. J., D. Wang, and D. J. Verschuur, 2008b, Adaptive curvelet-domain primary-multiple separation: Geophysics, **73**, no. 3, A17–A21.

Paige, C. C., and M. A. Saunders, 1982, LSQR, An algorithm for sparse linear equations and sparse least squares: ACM Trans. Math. Software, Volume 8, p. 43-71, **8**, 43–71.

Saab, R., D. Wang, O. Yilmaz, and F. Herrmann, 2007, Curvelet-based primary-multiple separation from a bayesian perspective: Presented at the SEG International Exposition and 77th Annual Meeting.

Vasconcelos, I., J. Gaiser, A. Calvert, and C. Calderón-Macías, 2008, Retrieval and suppression of surface waves using interferometry by correlation and by deconvolution: SEG Technical Program Expanded Abstracts, 2566–2570, SEG.

Verschuur, D. J., and A. J. Berkhout, 1997, Estimation of multiple scattering by iterative inversion, part II: practical aspects and examples: Geophysics, **62**, 1596–1611.

Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, 1166–1177.

Vogel, C., 2002, Computational Methods for Inverse Problems: SIAM.

Wang, D., R. Saab, O. Yilmaz, and F. J. Herrmann, 2008, Bayesian wavefield separation by transform-domain sparsity promotion: Geophysics, **73**, no. 5.

Wapenaar, K., and J. Fokkema, 2006, Green's function representations for seismic interferometry: Geophysics, **71**, SI33–SI46.

Yarham, C., and F. J. Herrmann, 2008, Bayesian ground-roll seperation by curvelet-domain sparsity promotion: SEG Technical Program Expanded Abstracts, 2576–2580, SEG.