

Decoupling of elastic parameters with iterative linearized inversion

Denis Anikiev*, Boris Kashtan, Saint Petersburg State University, and Wim A. Mulder, Shell Global Solutions International BV & Delft University of Technology

SUMMARY

Three model parameters as a function of position describe wave propagation in an isotropic elastic medium. Ideally, imaging of data for a point scatterer that consists of a perturbation in one of the elastic parameters should only provide a reconstruction of that perturbation, without cross-talk into the other parameters. This is not the case for seismic migration, where a perturbation of one elastic parameter contributes to the images of all three model parameters. For a reliable reconstruction of the true elastic reflectivity, one can apply iterative migration or linearized inversion, where the misfit cost function is minimized by the conjugate-gradient method. We investigated the decoupling of the three isotropic elastic medium parameters with the iterative linearized approach. Instead of iterating, the final result can be obtained directly by means of Newton's method, using the pseudo-inverse of the Hessian matrix. Although the calculation of the Hessian for a realistic model is an extremely resource-intensive problem, it is feasible for the simple case of a point scatterer in a homogeneous medium, for which we present numerical results. We consider the iterative approach with the conjugate-gradient method and Newton's method with the complete Hessian. Experiments show that in both cases the elastic parameters are decoupled much better when compared to migration. The iterative approach achieves acceptable inversion results but requires a large number of iterations. For faster convergence, preconditioning is required. An optimal preconditioner, if found, can be used in other iterative methods including L-BFGS. We considered two well known types of preconditioners, based on diagonal and on block-diagonal Hessian approximations. Somewhat to our surprise, both preconditioners fail to improve the convergence rate. Hence, a more sophisticated preconditioning is required.

INTRODUCTION

Classic migration (Claerbout, 1971), as well as the sensitivity kernel (Liu and Tromp, 2008) by itself, can provide maps of the reflection coefficients, but it cannot determine the correct amplitudes of reflectors (Zhu et al., 2009). An isotropic elastic medium can be characterized by three parameters, e.g., density ρ , compressional-wave velocity α and shear-wave velocity β . The migration result carries information about inhomogeneities in each elastic parameter, but a perturbation of one of the parameters will appear in the migration image as a perturbation of all three parameters. Thus, the elastic parameters are coupled in the context of the solution of the inverse problem. For a reliable reconstruction of the true reflectivity, iterative migration or linearized inversion can be applied (Beydoun and Mendes, 1989; Jin et al., 1992; Tura and Johnson, 1993, a.o.). Østmo et al. (2002) described the advantages of the linearized

approach over the full non-linear problem when applying iterative migration to the acoustic constant-density wave equation.

Here, we investigate the decoupling of the elastic parameters in the context of iterative linearized inversion. A brief review of the theory is followed by a numerical study for the case of a point scatterer in a homogeneous isotropic elastic background model. The numerical results provide insight into the best attainable resolution and reveal the maximum amount of decoupling of the elastic model parameters that one can reach with the linearized inversion approach.

THEORY

We start with the frequency-domain equations of motion in an isotropic elastic medium, written as $\mathbf{L}\mathbf{u} = \mathbf{f}$. With $\mathbf{L}\mathbf{u}$ being the elastic wave operator $\mathbf{L}$ acting on the displacement vector $\mathbf{u}$ we have

$$\mathbf{L}\mathbf{u} = -\omega^2 \rho \mathbf{u} - \nabla \cdot [\lambda (\nabla \cdot \mathbf{u}) \mathbf{I} + \mu (\nabla \mathbf{u} + [\nabla \mathbf{u}]^T)] = \mathbf{f}. \quad (1)$$

In equation 1, the source term is expressed by $\mathbf{f}$, ρ is the mass density, λ and μ are the Lamé parameters, ω is the angular frequency and $\mathbf{I}$ is the identity tensor. For imaging, a formal elastic parameter m can be treated as a sum of a known background value m_0 and a perturbation m_s : $m = m_0 + m_s$. Linearization with the Born approximation produces two equations, $L_0 \mathbf{u}_0 = \mathbf{f}$ and $L_0 \mathbf{u}_s = -L_s \mathbf{u}_0$, where $\mathbf{u}_0$ is the incident wave field due to a source in the background model, $\mathbf{u}_s = \mathbf{u} - \mathbf{u}_0$ is a wave field scattered by perturbations, L_0 is the elastic wave operator with background parameters ρ_0 , λ_0 , μ_0 and L_s is the operator with the perturbations ρ_s , λ_s and μ_s . Instead of perturbation parameter m_s , it is convenient to use reflectivity given by

$$\tilde{m} = \frac{m}{m_0} - 1 = \frac{m_s}{m_0}. \quad (2)$$

The inverse problem consists in finding an optimal reflectivity model $\tilde{m}_{opt}$ among all models $\tilde{m}$ for given background parameters $\mathbf{m}_0$ and an observed scattered wave field $\mathbf{u}_{obs}$ by minimization of the least-squares misfit functional

$$J(\tilde{m}) = \frac{1}{2} \sum_{\omega} \sum_{\mathbf{x}^s, \mathbf{x}^r} \|\mathbf{u} - \mathbf{u}_{obs}\|^2. \quad (3)$$

The optimal reflectivity model corresponds to the minimum of the functional J and hence to the zero of the gradient of J with respect to the model parameters. The linear approximation of the gradient at $\tilde{m}_{opt}$ around a nearby initial reflectivity model leads to Newton's method (Fichtner, 2010),

$$H \tilde{m}_{opt} = -\nabla J, \quad (4)$$

where ∇J is the gradient of J with respect to the model parameters and H is the Hessian, which is the matrix of second derivatives of J with respect to those parameters. The minimization of the misfit functional J defined by equation 3 with

Decoupling of elastic parameters with iterative linearized inversion

Newton's method requires the computation and inversion of the Hessian matrix. The Hessian corrects for acquisition imprint and geometrical spreading. However, the computation of the complete Hessian matrix is out of reach for large-scale applications, although it is feasible for small-sized problems (Gélis et al., 2007).

Alternatively, the conjugate-gradient method (CGM) can be applied for the solution of equation 4. CGM avoids the explicit computation and storing of the Hessian, but only involves a matrix-vector product of the Hessian with the current solution at every iteration (Axelsson, 1996, e.g.). This matrix-vector product is just the gradient. A finite number of iterations with CGM should provide a result close to the one obtained by Newton's method and the pseudo-inverse of the Hessian.

If the number of iterations is large, a preconditioning can speed up convergence (Axelsson, 1996). But convergence speed is not the only important aspect of preconditioning. Obviously, if the pseudo-inverse of Hessian is used as preconditioner, the solution will be obtained in a single iteration and the method will be analogous to the Newton's method. In that case, there is no need for an iterative scheme. The preconditioner should be an effective approximation of the Hessian's inverse with the property that it brings the total computational cost of the iterations below that of the full Hessian. This involves the cost of its construction, its application at each iteration and the total number of iteration steps.

One commonly used and easily implemented preconditioner is Jacobi, constructed from the inverse of the diagonal part of the Hessian (Kelley, 1995, e.g.). Another obvious choice for a preconditioner is the inverse of the block-diagonal approximation of the Hessian (Beylkin and Burrige, 1990). This approximation neglects interactions between neighboring points in the Hessian and take into account only the cross-coupling between multiple model parameters in one spatial point (Beydoun and Mendes, 1989). Each block of the preconditioner is the inverse of 3×3 matrix with elements corresponding to the second derivatives of the misfit functional computed in one spatial point. If the simple preconditioners are inefficient, a more complicated preconditioning strategy will probably be necessary. Examples are ILU, incomplete Cholesky and approximate inverse factorization methods (Saad and van der Vorst, 2000, e.g.).

Full Waveform Inversion (FWI) techniques use the nonlinear conjugate-gradient method for minimization of misfit functional, or more recently, the BFGS quasi-Newton method (Mulder and Plessix, 2004; Métivier et al., 2012, a.o.). The BFGS or its limited-memory version L-BFGS method is also applicable to the solution of a linear algebraic system such as equation 4. Nazareth (1979) showed that the preconditioned CGM is a special case of the BFGS method. Moreover, the BFGS and L-BFGS methods are sensitive to the initial inverse Hessian, which can be viewed as a preconditioner. Thus, any conclusion on optimal preconditioning strategy for CGM can be easily generalized to the BFGS method.

NUMERICAL RESULTS

We considered the simple model of a point scatterer in a homogeneous isotropic elastic background, allowing us to compute the complete Hessian matrix explicitly and to study the decoupling of the elastic parameters. The background medium had a density ρ of 2 g/cm^3 , a P-wave velocity α of 2 km/s , and an S-wave velocity β of 1.2 km/s . Three types of point scatterer located at $x_p = 0 \text{ m}$, $y_p = 0 \text{ m}$ and $z_p = 750 \text{ m}$ have been considered: first with perturbation in density ρ , second with perturbation in P-wave impedance $Z_\alpha = \rho\alpha$ and third with perturbation in S-wave impedance $Z_\beta = \rho\beta$. A line of 152 shots was placed on the surface in the same plane $y = 0 \text{ m}$ as the scatterer. Therefore, only P- and SV-waves were involved. Shots were located along the x -axis between -1887.5 m and 1887.5 m at a 25-m interval. 153 receivers were also deployed on the surface along the x -axis between -1900 m and 1900 m at the same interval. We restricted ourselves to the case of a vertical-force source and vertical-component data.

The scattered wave field for a given frequency can be constructed from the 3-D Green functions in a homogeneous background (Wu and Aki, 1985, e.g.). We used a Ricker wavelet with a peak frequency of 15 Hz and 166 discrete frequencies from 0 to 42 Hz to compute the wave fields. To estimate the unknown reflectivity around the actual position of the scatterer, we took a square zone of 400 m by 400 m with a grid spacing of 10 m . Figure 1 displays the components

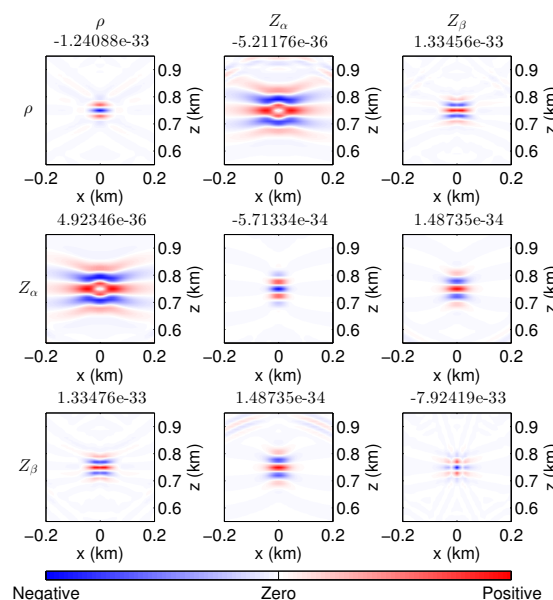


Figure 1: Components of the gradient of the misfit functional for three model cases. The column labels at the top shows which elastic parameter was perturbed. The row labels on the left indicate the component of the gradient, i.e., the parameter of differentiation. The images are scaled to their largest absolute value, the signed value of which is given on top of each picture. Decoupling is poor.

Decoupling of elastic parameters with iterative linearized inversion

of the gradient of the misfit functional. Each component of the gradient represents the derivative of the misfit functional with respect to one elastic parameter in a certain point of the medium. It determines if and by how much the corresponding element of the reflectivity model should be updated. In Figure 1, the three columns correspond to three cases of initial reflectivity at the scatterer: $\tilde{\rho}(\mathbf{x}) = A\delta(\mathbf{x} - \mathbf{x}_p)$, $\tilde{Z}_\alpha(\mathbf{x}) = 0$ and $\tilde{Z}_\beta(\mathbf{x}) = 0$ for the first column, $\tilde{\rho}(\mathbf{x}) = 0$, $\tilde{Z}_\alpha(\mathbf{x}) = A\delta(\mathbf{x} - \mathbf{x}_p)$ and $\tilde{Z}_\beta(\mathbf{x}) = 0$ for the second column, and $\tilde{\rho}(\mathbf{x}) = 0$, $\tilde{Z}_\alpha(\mathbf{x}) = 0$ and $\tilde{Z}_\beta(\mathbf{x}) = A\delta(\mathbf{x} - \mathbf{x}_p)$ for the third. The amplitude A was chosen such that we obtained a unit amplitude after representation on a grid. The rows correspond to the three components of the reflectivity, i.e., the derivatives with respect to $\tilde{\rho}$, $\tilde{Z}_\alpha$ and $\tilde{Z}_\beta$. The images have been scaled to the amplitude of the corresponding gradient component and the signed value that determined the maximum absolute value is given on top of each image. The color bar helps to distinguish between negative (blue) and positive (red) values of the gradient. Non-zero amplitudes are concentrated in the central area of the model, where the scatterer is located. However, all three elastic parameters appear to be sensitive to the perturbation, independently of its type. This implies strong coupling of these parameters. Therefore, the gradient alone is not enough for effective decoupling of the elastic parameters. In addition, the P- and S-waves are mixed in the gradient image and produce artefacts (c.f., Hak and Mulder, 2007). Figure 2 demonstrates that the coupling of

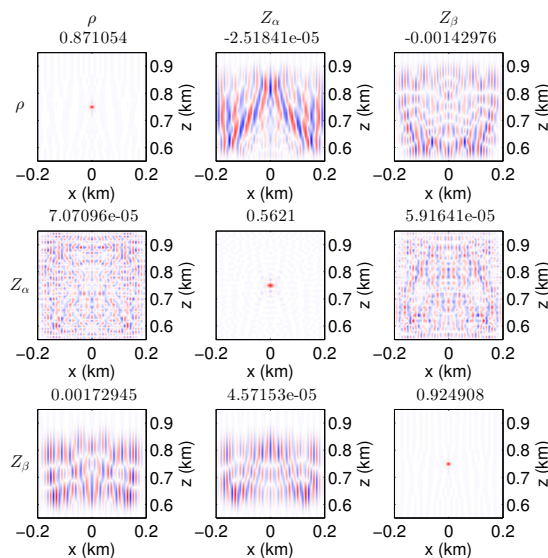


Figure 2: Reflectivities $\tilde{\rho}$, $\tilde{Z}_\alpha$ and $\tilde{Z}_\beta$ reconstructed by using the pseudo-inverse of the full Hessian. The column labels show which elastic parameter was perturbed. The row labels indicate the elastic parameter for which the reflectivity was reconstructed. These figures represent the best result that an iterative method can yield.

elastic parameters is much weaker when the pseudo-inverse of the Hessian is applied. The column labels again correspond to the elastic parameter that had its reflectivity perturbed at the lo-

cation of the point scatterer and the rows correspond to the reconstructed reflectivities $\tilde{\rho}$, $\tilde{Z}_\alpha$ and $\tilde{Z}_\beta$, respectively. The pictures are scaled by their amplitudes and the same color scheme as in Figure 1 is used. Ideal decoupling would result in the diagonal images having an amplitude 1 and the off-diagonal images having amplitudes equal to zero. In order to evaluate the decoupling of elastic parameters we introduced the quality of decoupling as

$$Q_a = \frac{|\tilde{M}_a| - \max(|\tilde{M}_b|, |\tilde{M}_c|)}{|\tilde{M}_a|} \cdot 100\%, \quad (5)$$

where $\tilde{M}_a$ is a maximum reflectivity for perturbed parameter a and $\tilde{M}_b$, $\tilde{M}_c$ are maximum reflectivities of two remaining unperturbed parameters b and c . Figure 2 indicates that $\tilde{\rho}$ and $\tilde{Z}_\beta$ are slightly coupled, whereas $\tilde{Z}_\alpha$ is coupled far less to the others. In terms of quality of decoupling, this corresponds to $Q_\rho \approx 99.8\%$, $Q_{Z_\alpha} \approx 100\%$ and $Q_{Z_\beta} \approx 99.8\%$.

CGM usually requires a large number of iterations. According to Kaporin (2012), this number can be roughly estimated as $n \leq \log_2 K$, where K is so-called K-condition number, which is a good alternative to the standard spectral condition number (Axelsson, 1996). In our case this bound approximately equals 60000. However, one may stop iterations earlier, when a reasonable solution quality is obtained. Therefore, we set the iterative scheme to stop when the quality of decoupling, defined in equation 5, was better than 95%. Figure 3 shows the

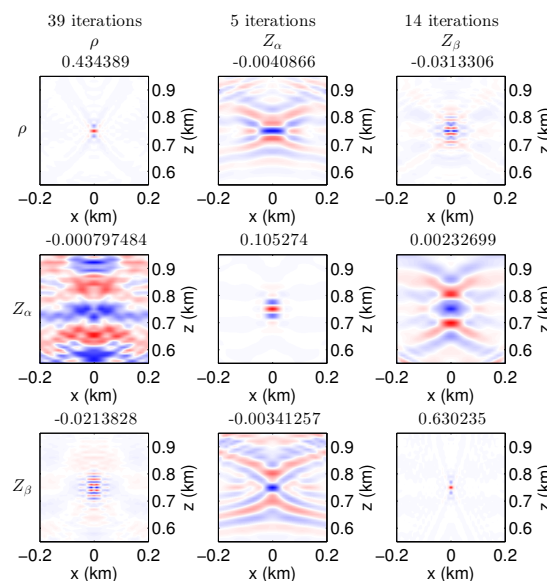


Figure 3: Reflectivities $\tilde{\rho}$, $\tilde{Z}_\alpha$ and $\tilde{Z}_\beta$ reconstructed by using CGM. The row and column labels are the same as in the earlier figures. The number of applied iterations is shown at the top of each column.

reconstruction of the reflectivities after using CGM. The number of iterations was different for each column and is given at the top. Thus CGM required 39 iterations to reach the desired quality for a perturbation in the density reflectivity, 5 itera-

Decoupling of elastic parameters with iterative linearized inversion

tions for a perturbation in $\tilde{Z}_\alpha$ and 14 iterations for $\tilde{Z}_\beta$. As already mentioned, the quality of decoupling in each case is better than 95%. Therefore, the iterative approach is able to give a satisfactory quality of decoupling, provided that number of iterations is sufficient. However, without a quality condition in mind, a standard tolerance of 10^{-6} can be reached within 8,385 iterations for $\tilde{\rho}$, 12,278 iterations for $\tilde{Z}_\alpha$ and 3,381 iterations for $\tilde{Z}_\beta$, which agrees with the derived upper bound. But even 39 iterations of CGM may be too large for realistic problems. Then, a suitable preconditioner may improve the convergence rate. Figure 4 shows the result of inversion by

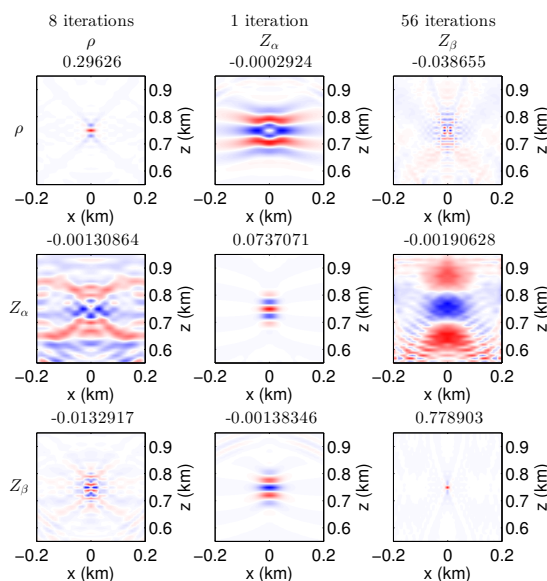


Figure 4: Reflectivities $\tilde{\rho}$, $\tilde{Z}_\alpha$ and $\tilde{Z}_\beta$ reconstructed using CGM preconditioned by the inverse of the diagonal part of the Hessian. The row and column labels are the same as in the earlier figures.

CGM preconditioned by the inverse diagonal part of Hessian (not the the block-diagonal part). This simple preconditioning decreased the number of iterations to 8 in case of $\tilde{\rho}$ and to 1 in case of $\tilde{Z}_\alpha$, but for $\tilde{Z}_\beta$, it increased the number of iteration to 56. Therefore, depending on the problem, Jacobi preconditioner can either speed up convergence or slow it down. Result in the case when a more expensive block-diagonal preconditioning was used is shown in Figure 5. Although the number of iterations decreased to 6 for $\tilde{\rho}$, an extremely large number of 3,828 iterations resulted for $\tilde{Z}_\beta$. For the third column in Figure 5, the decoupling quality is around 88%, i.e., the result has insufficient quality. Therefore, an inverse block-diagonal preconditioner is not the best choice for proper decoupling.

CONCLUSIONS

We investigated the decoupling of the three isotropic elastic medium parameters in the context of iterative linearized inversion. Numerical inversion experiments were presented for

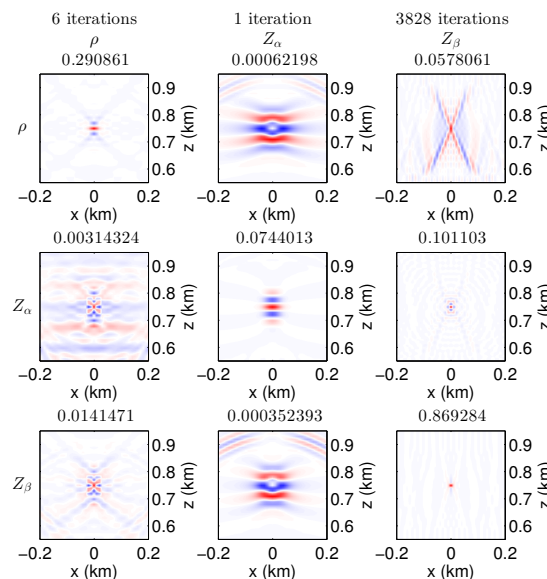


Figure 5: Reflectivities $\tilde{\rho}$, $\tilde{Z}_\alpha$ and $\tilde{Z}_\beta$ reconstructed with CGM preconditioned by the inverse of the block-diagonal part of the Hessian. The row and column labels are the same as in the earlier figures.

Newton's method, using the pseudo-inverse of Hessian matrix, and for the conjugate-gradient method with different preconditioners. We conclude that with both methods, the elastic parameters are decoupled much better when compared to migration. The standard iterative approach allows us to achieve acceptable inversion results but requires a large number of iterations. Preconditioning can help to speed up convergence. Common wisdom suggests using the inverse of diagonal or block-diagonal part of the Hessian as a preconditioner. In our examples we show that such a choice is far from optimal. An optimal preconditioning should possess several properties simultaneously: it must be relatively close to the inverse of the Hessian matrix; the cost of computing and applying the preconditioner at every iteration must be relatively low; preconditioning must allow for efficient parallelization and, finally, it must improve the convergence rate.

In general, our intention is to find an appropriate imaging conditions which allow for the best decoupling of the isotropic elastic scattering parameters. These conditions involve not only a suitable preconditioner for an iterative method, but problem parametrization issues, wave field decomposition, Hessian SVD analysis and other theoretical considerations.

ACKNOWLEDGMENTS

The authors are grateful to the team of Elastic Media Dynamics Laboratory of Saint Petersburg State University for helpful discussions. The research is supported by Shell Global Solutions International BV through CRDF grant RUG1-30018-ST-11.

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2013 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

- Axelsson, O., 1996, Iterative solution methods: Cambridge University Press.
- Beydoun, W. B., and M. Mendes, 1989, Elastic ray-born L2-migration/inversion: *Geophysical Journal International*, **97**, 151–160, <http://dx.doi.org/10.1111/j.1365-246X.1989.tb00490.x>.
- Beylkin, G., and R. Burridge, 1990, Linearized inverse scattering problems in acoustics and elasticity: *Wave Motion*, **12**, 15–52, [http://dx.doi.org/10.1016/0165-2125\(90\)90017-X](http://dx.doi.org/10.1016/0165-2125(90)90017-X).
- Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: *Geophysics*, **36**, 467–481, <http://dx.doi.org/10.1190/1.1440185>.
- Fichtner, A., 2010, Full seismic waveform modelling and inversion: Springer Verlag.
- Gélis, C., J. Virieux, and G. Grandjean, 2007, Two-dimensional elastic full waveform inversion using Born and Rytov formulations in the frequency domain: *Geophysical Journal International*, **168**, 605–633, <http://dx.doi.org/10.1111/j.1365-246X.2006.03135.x>.
- Hak, B., and W. A. Mulder, 2007, A remark on the seismic resolution function for the homogeneous elastic isotropic case: 77th Annual International Meeting, SEG, Expanded Abstracts, 2422–2426, <http://dx.doi.org/10.1190/1.2792970.x>.
- Jin, S., R. Madariaga, J. Virieux, and G. Lambaré, 1992, Two-dimensional asymptotic iterative elastic inversion: *Geophysical Journal International*, **108**, 575–588, <http://dx.doi.org/10.1111/j.1365-246X.1992.tb04637.x>.
- Kaporin, I. E., 2012, Using Chebyshev polynomials and approximate inverse triangular factorizations for preconditioning the conjugate gradient method: *Computational Mathematics and Mathematical Physics*, **52**, 169–193, <http://dx.doi.org/10.1134/S0965542512020091>.
- Kelley, C., 1995, Iterative methods for linear and nonlinear equations: SIAM.
- Liu, Q., and J. Tromp, 2008, Finite-frequency sensitivity kernels for global seismic wave propagation based upon adjoint methods: *Geophysical Journal International*, **174**, 265–286, <http://dx.doi.org/10.1111/j.1365-246X.2008.03798.x>.
- Métivier, L., R. Brossier, J. Virieux, and S. Operto, 2012, The truncated Newton method for full waveform inversion: 82nd Annual International Meeting, SEG Expanded Abstracts, <http://dx.doi.org/10.1190/segam2012-0981.1>.
- Mulder, W. A., and R.-E. Plessix, 2004, A comparison between one-way and two-way wave-equation migration: *Geophysics*, **69**, 1491–1504, <http://dx.doi.org/10.1190/1.1836822>.
- Nazareth, L., 1979, A Relationship between the BFGS and conjugate gradient algorithms and its implications for new algorithms: *SIAM Journal on Numerical Analysis*, **16**, 794–800, <http://dx.doi.org/10.1137/0716059>.

- Østmo, S., W. A. Mulder, and R.-E. Plessix, 2002, Finite-difference iterative migration by linearized waveform inversion in the frequency domain: 82nd Annual International Meeting, SEG, Expanded Abstracts, 1384–1387, <http://dx.doi.org/10.1190/1.1816917>.
- Saad, Y., and H. A. van der Vorst, 2000, Iterative solution of linear systems in the 20th century: *Journal of Computational and Applied Mathematics*, **123**, 1–33, [http://dx.doi.org/10.1016/S0377-0427\(00\)00412-X](http://dx.doi.org/10.1016/S0377-0427(00)00412-X).
- Tura, M. A. C., and L. R. Johnson, 1993, A stable method for linearized inversion of elastic parameters: *Geophysical Journal International*, **115**, 1–13, <http://dx.doi.org/10.1111/j.1365-246X.1993.tb05585.x>.
- Wu, R., and K. Aki, 1985, Scattering characteristics of elastic waves by an elastic heterogeneity: *Geophysics*, **50**, 582–595, <http://dx.doi.org/10.1190/1.1441934>.
- Zhu, H., Y. Luo, T. Nissen-Meyer, C. Morency, and J. Tromp, 2009, Elastic imaging and time-lapse migration based on adjoint methods: *Geophysics*, **74**, no. 6, WCA167–WCA177, <http://dx.doi.org/10.1190/1.3261747>.