Nonlinear Fréchet derivative and its De Wolf approximation

Ru-Shan Wu[1] and Yingcai Zheng[2]***
[1] Modeling and Imaging Laboratory, University of California, Santa Cruz, CA
[2]Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA

Summary

We introduce and derive the nonlinear Fréchet derivative for the acoustic wave equation. It turns out that the high order Fréchet derivatives can be realized by consecutive applications of the scattering operator and a zero-order propagator to the source. We prove that the higher order Fréchet derivatives are not negligible and the linear Fréchet derivative may not be appropriate in many cases, especially when forward scattering is involved for large scale perturbations. Then we derive the De Wolf approximation (multiple forescattering and single backscattering approximation) for the nonlinear Fréchet derivative. We split the linear derivative operator (i.e. the scattering operator) onto forward and backward derivatives, and then reorder and renormalize the nonlinear derivative series before making the approximation by dropping the multiple backscattering terms. Numerical simulations for a Gaussian ball model show significant difference between the linear and nonlinear Fréchet derivatives.

Introduction

The Fréchet derivative plays a key role in linear or quasi-linear geophysical inverse problems. The Fréchet derivative is usually referred to the first order derivative. Higher order terms are thought to be insignificant or of no importance, and are generally neglected. In seismic traveltime tomography, explicit formula for the Fréchet derivative has been derived , and numerical calculations have been carried out (e.g., Marquering, *et al*., 1999; Dahlen *et al.*, 2000a; Dahlen *et al.*, 2000b; Hung *et al.*, 2000; Dahlen, 2004; 2005; Dahlen and Nolet, 2006; de Hoop and van der Hilst, 2005; Zhao et al., 2000; Zhou et al., 2004, 2011). This spatial distribution of the traveltime changes caused by the velocity perturbations at the corresponding points is called sensitivity kernel for traveltimes. The derivation of the kernels is usually calculated by the first order Born approximation based on the formulation of linearized Fréchet derivative. Similar sensitivity kernel formulation has been used in reflection seismics (de Hoop et al., 2006; Xie and Yang, 2008). For the full waveform inversion, the Fréchet derivative is also used to relate model parameter perturbations to changes in seismic waveforms (Tarantola, 1984; Tarantola, 1986; Tarantola, 2005). In order to mitigate the problem of huge storage for the kernels based on the point scattering model, Chevrot and Zhao (2007) proposed to use wavelet transform applied to the model space; Loris et al. (2010) and Simon et al. (2011) apply wavelet transform to the model for the purpose of L-1 norm regularization in the inversion. However, in both cases the linearized Fréchet derivatives are used for the sensitivity kernel calculations.

For the real earth, the wave equation is strongly nonlinear with respect to the medium parameter changes, except in some weakly perturbed media. For large scale velocity perturbations, the phase accumulation of forward scattering renders the Born approximation unacceptable in many cases. However, based on our knowledge, the kernel calculations are almost exclusively based on the linear Fréchet derivative in the literature. Nonlinear Fréchet derivative has been introduced to the resistivity inversion (McGillivray and Oldenburg, 1990) and optical diffuse imaging (Kwon and Yazici, 2010). In this study, we first discuss the sensitivity kernel for the acoustic wave equation and then we derive the full nonlinear Fréchet derivative as a series with all the higher order terms. In the second part, we derive the the De Wolf approximation of the nonlinear Fréchet derivative. A numerical experiment is conducted for a Gaussian ball with different scales. The results show significant differences between the linear and nonlinear Fréchet derivatives.

Linear and Nonlinear Fréchet derivatives for the Acoustic Wave Equation

For a linear isotropic acoustic medium, the wave equation in frequency domain is

$$\nabla \cdot \frac{1}{\rho} \nabla p + \frac{\omega^2}{\kappa} p = 0 \tag{1}$$

where p is the pressure field, ρ and κ are the density and bulk module of the medium, respectively. Assuming ρ_0 and κ_0 as the parameters of the background medium, equation (1) can be written as

$$\left(\nabla^2 + k^2\right) p(\mathbf{x}) = -k^2 \varepsilon(\mathbf{x})\, p(\mathbf{x}), \tag{2}$$

where

$$k = \omega / v_0; \qquad v = \sqrt{\kappa_0 / \rho_0} \ . \tag{3}$$

The right-hand-side of (2) is an equivalent force term with

$$\varepsilon(\mathbf{x}) = \varepsilon_\kappa(\mathbf{x}) + \frac{1}{k^2} \nabla \cdot \varepsilon_\rho \nabla \tag{4}$$

as the scattering potential operator, where

$$\varepsilon_\rho(\mathbf{x}) = \frac{\rho_0}{\rho(\mathbf{x})} - 1, \quad \varepsilon_\kappa(\mathbf{x}) = \frac{\kappa_0}{\kappa(\mathbf{x})} - 1, \tag{5}$$

Equation (2) can be written into an equivalent integral equation form (Lippmann-Schwinger equation):

$$p(\mathbf{x}) = p^0(\mathbf{x}) + k^2 \int_V d^3\mathbf{x}' g_0(\mathbf{x};\mathbf{x}') \varepsilon(\mathbf{x}') p(\mathbf{x}') \qquad (6)$$

where $g_0(\mathbf{x};\mathbf{x}')$ is the background Green's function. In operator form, equation (6) can be written as

$$p = p^0 + G_0 \varepsilon\, p \qquad (7)$$

The above quation is expressed as a summation of the incident field and the perturbed field (scattered field). Substitute this sum into the unknown p in the right-hand-side iteratively, resulting in a formal solution of the wave equation in a scattering series (the Born series)

$$\begin{aligned} p &= p^0 + G_0\varepsilon\, p^0 + G_0\varepsilon\, G_0\varepsilon\, p^0 + \cdots\cdots \\ &= \sum_{n=0}^{\infty} [G_0\varepsilon]^n\, p^0 \end{aligned} \qquad (8)$$

Now we consider the Fréchet derivatives for the forward modeling operator. In accordance with Tarantola (2005, Chapter 5)'s format and terminology, we write the forward problem in an operator form,

$$\mathbf{d} = \mathbf{A}(\mathbf{m}) \qquad (9)$$

where **d** is the data vector (pressure field generated by the modeling), **m** is the model vector, and **A** is the forward modeling operator. Assume an initial model $\mathbf{m_0}$, we want to quantify the sensitivity of the data change $\delta\mathbf{d}$ to the model change $\delta\mathbf{m}$. For a linear modeling operator, or a quasi-linear operator, we can calculate the data change using the linear Fréchet derivative (Tarantola, 2005)

$$\mathbf{A}(\mathbf{m}_0 + \delta\mathbf{m}) = \mathbf{A}(\mathbf{m}_0) + \mathbf{A}'\delta\mathbf{m} + O(\|\delta\mathbf{m}\|^2) \qquad (10)$$

where $\mathbf{A}'$ is the first Fréchet derivative operator ("G_0" in Tarantola's notation). Comparing (10) with (7), we see that the linear Fréchet derivative operator is equivalent to a Born modeling operator

$$\mathbf{A}' = G_0\ \varepsilon p^0 = k^2 \int_V d^3\mathbf{x}' g_0(\mathbf{x};\mathbf{x}') \varepsilon(\mathbf{x}') p^0(\mathbf{x}') \qquad (11)$$

The Higher order Fréchet derivative can be derived from the first order one. $\mathbf{A}'$ is a bounded linear operator which maps the model space $\mathcal{M}$ to the data space $\mathcal{D}$; The mathematical definition of the linear Fréchet derivative reads

$$\lim_{\delta\mathbf{m}\to 0} \frac{\|A(\mathbf{m}_0 + \delta\mathbf{m}) - A(\mathbf{m}_0) - A'(\mathbf{m}_0)\delta\mathbf{m}\|_{\mathcal{D}}}{\|\delta\mathbf{m}\|_{\mathcal{M}}} = 0 \qquad (12)$$

We define the space of all such bounded operators as

$$L(\mathcal{M},\mathcal{D}) = \{A'(\mathbf{m}_0) \mid \mathbf{m}_0 \in \mathcal{M}\}. \qquad (13)$$

Higher order Fréchet derivatives can be derived from the first order Fréchet derivative. Let's first see the second order derivative

$$\lim_{\delta\mathbf{m}_2\to 0} \frac{\|A'(\mathbf{m}_0 + \delta\mathbf{m}_2) - A'(\mathbf{m}_0) - A''(\mathbf{m}_0)\delta\mathbf{m}_2\|_{L(\mathcal{M},\mathcal{D})}}{\|\delta\mathbf{m}_2\|_{\mathcal{M}}} = 0 \qquad (14)$$

We plug (11) into (14) and after some algebra we get

$$\begin{aligned} A''(\mathbf{m}_0)\delta\mathbf{m}^2 &= 2\int d^3\mathbf{x}' k_0^2 g_0(\mathbf{x} \mid \mathbf{x}')\delta m(\mathbf{x}') \\ &\times \int d^3\mathbf{x}'' k_0^2 g_0(\mathbf{x}' \mid \mathbf{x}'')\, p^0(\mathbf{x}'')\delta m_2(\mathbf{x}'') \end{aligned} \qquad (15)$$

Other higher orders can be derived in a similar way. Put all orders together, we have a Taylor expansion centered at the original model

$$\begin{aligned} &\mathbf{A}(\mathbf{m}_0 + \delta\mathbf{m}) = \mathbf{A}(\mathbf{m}_0) + \mathbf{A}'(\mathbf{m}_0)\delta\mathbf{m} \\ &+ \frac{1}{2!}\mathbf{A}''(\mathbf{m}_0)(\delta\mathbf{m})^2 + \ldots + \frac{1}{n!}\mathbf{A}^{(n)}(\mathbf{m}_0)(\delta\mathbf{m})^n \\ &+ \ldots\ldots = \mathbf{A}(\mathbf{m}_0) + \mathbf{A}_{\mathbf{m}_0}^{(NLD)}(\delta\mathbf{m}) \end{aligned} \qquad (16)$$

where $\mathbf{A}_{\mathbf{m}_0}^{(NLD)}$ is the nonlinear functional derivative at the current model. Under the linear approximation, $\mathbf{A}'\delta\mathbf{m}$ is a matrix multiplication in the discrete form, so that the influences of parameter changes at different points are independent of each other. In the case of wave equation, this linearization is only valid for weak heterogeneities and small volume perturbations. On the other hand, the nonlinear derivative (16), if the series converge, can precisely predict the data change $\delta\mathbf{d}$ due to the model change $\delta\mathbf{m}$. The higher order terms account for the interactions between different parameter changes (multiple scattering).

For the acoustic media, we apply the Gauss theory to the volume integral in (6). After some algebra, we get

$$\begin{aligned} &\mathbf{A}'(\mathbf{m}_0)\delta\mathbf{m} \\ &= k^2 \int_{V_1} d^3\mathbf{x}_1 \left\{ \left[g_0(\mathbf{x}_r;\mathbf{x}_1) g_0(\mathbf{x}_1;\mathbf{x}_s) \right] \varepsilon_\kappa(\mathbf{x}_1) \right. \\ &\left. + \left[\bar{\nabla}_1 g_0(\mathbf{x};\mathbf{x}_1) \cdot \bar{\nabla}_1 g_0(\mathbf{x}_1;\mathbf{x}_s) \right] \varepsilon_\rho(\mathbf{x}_1) \right\} \end{aligned} \qquad (17)$$

where $\bar{\nabla} \equiv \frac{1}{ik}\nabla$ is the frequency-normalized gradient operator, and the subscript 1 implies that the derivative is related to $\mathbf{x}_1$. In operator form, we can write the above equation as

$$\mathbf{A}'(\mathbf{m}_0)\delta\mathbf{m} = \left(\mathbf{S}_\kappa \mathbf{A}^0, \mathbf{S}_\rho \mathbf{A}^0 \right) \begin{pmatrix} \varepsilon_\kappa \\ \varepsilon_\rho \end{pmatrix} \qquad (18)$$

Where

$$\mathbf{A}^0 = \mathbf{A}(\mathbf{m}_0) = g_0(\mathbf{x}_1;\mathbf{x}_s) \qquad (19)$$

And $\mathbf{S}_\kappa, \mathbf{S}_\rho$ are the scattering operator defined as

$$\mathbf{S}_\kappa = k^2 \int_{V_1} d^3\mathbf{x}_1 g_0(\mathbf{x};\mathbf{x}_1)$$
$$\mathbf{S}_\rho = k^2 \int_{V_1} d^3\mathbf{x}_1 \overline{\nabla}_1 g_0(\mathbf{x};\mathbf{x}_1) \cdot \overline{\nabla}_1 \tag{20}$$

So the first Fréchet derivative is an operator

$$\mathbf{A}'(\mathbf{m}_0) = \mathbf{S}_{\kappa,\rho}\mathbf{A}^0 = (\mathbf{S}_\kappa, \mathbf{S}_\rho)\mathbf{A}^0 \tag{21}$$

We obtain the second order of Fréchet derivative operator as

$$\begin{aligned}\mathbf{A}''(\mathbf{m}_0) &= 2!(\mathbf{S}_\kappa, \mathbf{S}_\rho)^2 \mathbf{A}^0 \\ &= 2!(\mathbf{S}_\kappa\mathbf{S}_\kappa, \mathbf{S}_\kappa\mathbf{S}_\rho, \mathbf{S}_\rho\mathbf{S}_\kappa, \mathbf{S}_\rho\mathbf{S}_\rho)\mathbf{A}^0\end{aligned} \tag{22}$$

For the nth order Fréchet derivative, we can write as

$$\mathbf{A}^{(n)}(\mathbf{m}_0) = n!(\mathbf{S}_\kappa, \mathbf{S}_\rho)^n \mathbf{A}^0 \tag{23}$$

We see that the n^{th} order Fréchet derivative can be realized by consecutive application of the scattering operator:

$$\mathbf{A}^{(n)}(\mathbf{m}_0) = n!(\mathbf{S})^n \mathbf{A}^0 = n!(\mathbf{S}_n\mathbf{S}_{n-1}......\mathbf{S}_1)\mathbf{A}^0 \tag{24}$$

De Wolf Approximation of the Nonlinear Fréchet Derivative

The full nonlinear Fréchet derivative accounts for all the nonlinear interactions between the perturbations at different points with different parameters. To compute the full Fréchet derivative series is very time-consuming or intractable! In order to take the benefit provided by the Fréchet derivative series but keep the computation manageable, we turn to the De Wolf approximation for the full series computation. If we split the scattering operator into forward scattering (forescattering) and backscattering parts

$$\mathbf{S} = \mathbf{S}^f + \mathbf{S}^b \tag{25}$$

and substitute it into the Fréchet series, we can have all combinations of high order forward and backward derivatives. These high-order derivatives are similar to the high-order scattering terms in the Born series. The De Wolf approximation in scattering series corresponds to neglect the multiple backscattering (reverberations), i.e., drop all the terms containing two or more backscattering operators but keep all the forward scattering terms untouched (De Wolf, 1971, 1985; Wu, 1994; for a summary and review, see Wu et al., 2007). In the forward direction, scattered fields of $\delta\kappa$ and $\delta\rho$ have the opposite signs, so that the resulted scattering response is for the velocity perturbation; while in the backward direction, the scattered fields of $\delta\kappa$ and $\delta\rho$ have the same sign, which corresponds to a response of impedance perturbation. The goal of transmission tomography is to determine the velocity perturbations of the target structure and the physical process in the play is the forward scattering. Therefore, the multiple forward scattering is very important and the point-scatterer model based on the Born approximation may not be appropriate for transmission tomography.

After substituting the split scattering operator (25) into the expression of n^{th} order Fréchet derivative, we can obtain the general term of MFSB (multiple forescattering-single backscattering) approximation by sorting out and keep only the terms with one backscattering operator:

$$\mathbf{A}_{DW}^{(n)}(\mathbf{m}_0) = n!(\mathbf{S}_n^f\mathbf{S}_{n-1}^f...\mathbf{S}_{i+1}^f\mathbf{S}_i^b\mathbf{S}_{i-1}^f....\mathbf{S}_1^f)\mathbf{A}^0 \tag{26}$$

Finally, we can obtain the nonlinear Fréchet derivative (sensitivity kernel) as

$$\begin{aligned}&\mathbf{A}^{(NLD)}(\mathbf{x};\mathbf{x}_g,\mathbf{x}_s) = \\ &\quad \mathbf{S}_F(\mathbf{x}_g;\mathbf{x})\mathbf{S}_{\kappa,\rho}^b(\mathbf{x})\mathbf{S}_F(\mathbf{x};\mathbf{x}_1)\mathbf{A}^0(\mathbf{x}_1;\mathbf{x}_s)\end{aligned}, \tag{27}$$

where $\mathbf{S}_F$ is the multi-forescattering operator. Compare with the first order Fréchet derivative for reflection problem, we see the similarity in the form and the simplicity of the formulation. However, the Born scattering operator in (18) is replaced with the multi-forescattering operator. The nonlinear sensitivity kernel with the De wolf approximation considers all the nonlinear interactions with the point perturbations it passed under the forward scattering approximation, so that the phase accumulation and amplitude changes from the perturbations at other points are included.

Numerical Examples of Nonlinear Fréchet Derivative And Its De Wolf Approximation

In order to show the limitation of the linear Fréchet derivative and the merit of nonlinear Fréchet derivative with the De Wolf approximation (DWA), we conduct a set of forward scattering experiments using Gaussian shape velocity perturbations.

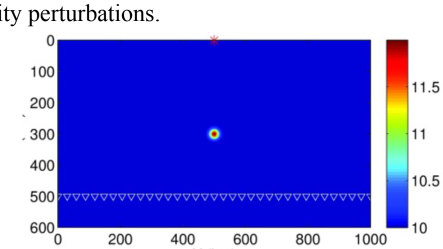


Figure 1: Experiment configuration similar to transmission tomography. The source is the red star (f_0=20 Hz Ricker) on the top. The receivers are white triangles near the bottom. The scale of the Gaussian anomaly is a= 20, 100 and 500 m, respectively.

As shown in Figure 1, the source is the red star on the top. The receivers are white triangles near the bottom. The velocity model is a fast Gaussian anomaly embedded in a constant background. The background velocity is 2 km/s. The scales of the Gaussian anomaly are a = 20*m*, 100*m* and

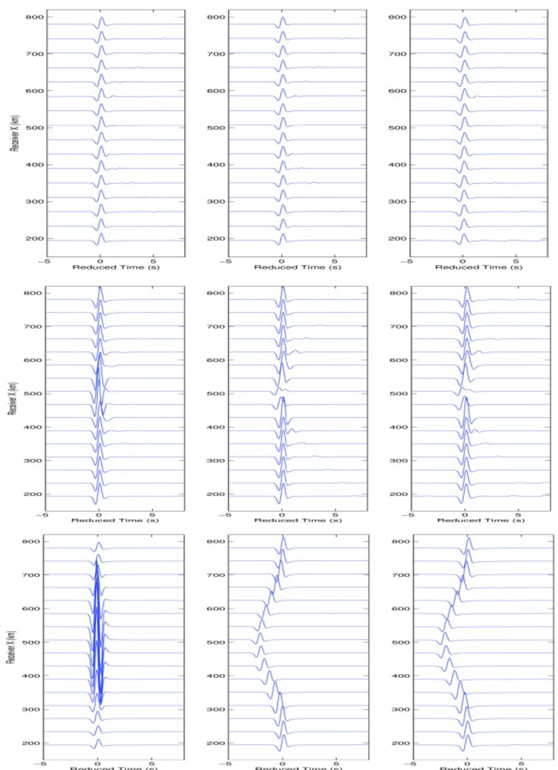
500*m*, respectively. A line of receivers are placed at $z = 5km$. For Born modeling, a simple summation integral is used (Aki, 1080); The FD modeler is a regular 4th order finite difference algorithm; The De Wolf modeling is realized by a GSP (generalized screen propagator) one-way propagator (see Wu, 2003), because there is no reflection involved. In the calculations, we include the evanescent wave components (inhomogeneous plane waves) (Aki, 1980) to reduce the wavenumber truncation artifacts). Figure 2 shows the waveform comparison from different kernels (Left: Born, i.e. Linear F-derivative; Mid: FD; Right: De Wolf approximation of NL F-derivative) for the Gaussian ball with a=20*m* (top), 100*m* (mid), 500*m* (bottom); We see from the figure that for the small scale perturbation, the Born approximation is more or less valid and we see the agreement between all the methods. However, for large scale perturbations, the deviations from Born approximation are significant. Especially in the case of a = 500*m*, the kernel predictions by linear Fréchet derivative become unacceptable. In the meanwhile, the forward scattering renormalized modeling, here the GSP propagator, can give fairly accurate waveforms. The striking large amplitude in the center shows the famous forward-scattering catastrophe of the Born approximation.

Conclusion and Discussion

We proved that the high order Fréchet derivatives can be realized by consecutive applications of the scattering operator and a zero-order propagator to the source. The full nonlinear Fréchet derivative is directly related to the full scattering series (Born series). We then split the linear derivative operator onto forward and backward derivatives, and then derive the De Wolf approximation (multiple forescattering and single backscattering approximation) for the full NL F-derivative. Through both theoretical analysis and numerical examples we demonstrated the inadequacy of the linear Fréchet derivative in many cases, especially when forward scattering is involved for large scale velocity perturbations. In those cases, the use of nonlinear F-derivative becomes important and the De Wolf approximation of NL F-derivative may be an efficient alternative for kernel calculation.

The linear F-derivative defines both the forward modeling (Born modeling) and the back projection through the gradient operator (Born inversion). Because of the linearization, the predicted model perturbations may be far from the real perturbations causing the data changes. This is true even when the perturbations are weak but extended for a large volume. This is why the inversion is strongly dependent on the initial model, especially on the low-wavenumber component of the model (smooth background). If we can handle the forward scattering correctly using the NL F-derivative, so that the backprojection operator can update the low-wavenumber components (smooth background) during the iteration, the dependence on the initial model may be much reduced.

Acknowledgments: The research is supported by the WTOPI (Wavelet Transform on Propagation and Imaging for seismic exploration) Research Consortium at the University of California, Santa Cruz and ERL/MIT.
Facility support from the W.M. Keck Foundation is also acknowledged.



Born FD De Wolf

Figure 2: Waveform comparison from different kernels for a fast Gaussian ball (20% at the center) with a=20 (top), 100 (mid), 500*m* (bottom); Left: Born (Linear F-derivative); Mid: FD; Right: De Wolf approximation of NL F-derivative. The reduced traveltime is defined as the true traveltime minus the background traveltime.

References

Aki, K., Richards, P. 1980, Quantitative Seismology: Theory and Methods, vols. 1 and 2. Freeman,New York.

Chevrot, S. and L. Zhao, 2007, Multiscale finite-frequency Rayleigh wave tomography of the Kaapvaal craton, *Geophys. J. Int*., 169, 201-215.

Dahlen, F. A. (2004), Resolution limit of traveltime tomography, *Geophys. J. Int*., 157(1), 315-331.

Dahlen, F. A. (2005), Finite-frequency sensitivity kernels for boundary topography perturbations, *Geophys. J. Int.*, 162(2), 525-540.

Dahlen, F. A., S. H. Hung, and G. Nolet, 2000a, Frechet kernels for finite-frequency traveltimes; I, Theory, *Geophys. J. Int.*, 141(1), 157-174.

Dahlen, F. A., S. H. Hung, G. Nolet, and Anonymous, 2000b, Banana-doughnut traveltime Frechet kernels, edited.

Dahlen, F.A. and Nolet, G., 2006, Comment on 'On sensitivity kernels for 'wave-equation' transmission tomography' by de Hoop and van der Hilst, *Geophys. J. Int*., 163, 949–951.

de Hoop, M. V., and R. D. van der Hilst, 2005, On sensitivity kernels for 'wave-equation' transmission tomography, *Geophys. J. Int*., 160(2), 621-633.

de Hoop, M. V., R. D. van der Hilst and S. Peng, 2006, Wave-equation reflection tomography: annihilators and sensitivity kernels, *Geophys. J. Int.,* **167,** 1332–1352.

De Wolf, D.A. (1971). Electromagnetic reflection from an extended turbulent medium: Cumulative forward-scatter single-backscatter approximation. *IEEE Trans. Antennas and Propagations* **AP-19**, 254–262.

De Wolf, D.A. (1985). Renormalization of EM fields in application to large-angle scattering from randomly continuous media and sparse particle distributions. *IEEE Trans. Antennas and Propagations***AP-33**, 608–615.

Hung, S. H., F. A. Dahlen, and G. Nolet (2000), Frechet kernels for finite-frequency traveltimes; II, Examples, *Geophys. J. Int*., 141(1), 175-203.

Kwon, K. and Yazici, B., 2010, Born expansion and Frechet derivatives in nonlinear diffuse optical tomography, *Computers and Mathematics with Applications*, doi: 10.1016/j.camwa.2009.07.088.

Loris, I., Douma, H., Nolet, G., Daubechies, I. & Regone, C., 2010, Nonlinear regularization techniques for seismic tomography, *J. Comput. Phys.*, 229(3), 890–905.

Marquering, H., Dahlen, F.A. & Nolet, G., 1999. Three-dimensional sensitivity kernels for ¢nite-frequency traveltimes: the banana-doughnut paradox, *Geophys. J. Int.,* **137**, 805-815.

McGillivray, P. R., and D. W. Oldenburg (1990), Methods for calculating Fréchet derivatives and sensitivities for the non-linear inverse problem: a comparative study, *Geophysical Prospecting*, 38(5), 499-524.

Nolet, G., 2008. *A Breviary for Seismic Tomography*, Cambridge Univ. Press, Cambridge, UK.

Simons, F.J., I. Loris, G. Nolet, I. C. Daubechies, S. Voronin, J. S. Judd, P. A. Vetter, J. Charlety and C. Vonesch, 2011, Solving or resolving global tomographic models with spherical wavelets, and the scale and sparsity of seismic heterogeneity, *Geophys. J. Int.,* **187**, 969-988.

Tarantola, A. (1984), Inversion of seismic-reflection data in the acoustic approximation, *Geophysics*, 49(8), 1259-1266.

Tarantola, A. (1986), A strategy for nonlinear elastic inversion of seismic reflection data, *Geophysics*, 51(10), 1893-1903.

Tarantola, A. (2005), Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, Philadelphia, PA.

Virieux, J. and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics, *Geophysics*, 74(6), wcc1-wcc26.

Wu, R.S., 1994, Wide-angle elastic wave one-way propagation in heterogeneous media and an elastic wave complex-screen method, *J. Geophys. Res.,* **99,** 751-766.

Wu, R.S., 2003, Wave propagation, scattering and imaging using dual-domain one-way and one-return propagators, *Pure and Appl. Geophys.*, 160(3/4), 509-539.

Wu, R.S., and K. Aki, 1985a, Scattering characteristics of waves by an elastic heterogeneity, *Geophysics,* 50, 582-595.

Wu, R.S**.**, Xie, X.B. and Wu, X.Y., 2007, One-way and one-return approximations for fast elastic wave modeling in complex media, Chapter 5 of " *Advances in Wave Propagation in Heterogeneous Earth*", edited by R.S. Wu and V. Maupin (V.48 of " *Advances in Geophysics*", Series editor R. Dmowska), Elsevier, 266-323.

Xie, X.B., and H., Yang, 2008, The finite-frequency sensitivity kernel for migration residual moveout and its applications in migration velocity analysis, *Geophysics*, 73, S241-S249.

Zhang, K. C., 2005, Methods in nonlinear analysis, Springer.

Zhao,L., T. H. Jordan and C. H. Chapman, 2000, Three-dimensional Frechet differential kernels for

seismic delay times, *Geophys. J. Int.,* 141, 558-576.
Zhou, Y., Dahlen, F.A. & Nolet, G., 2004. Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int*., 158, 142–168.
Zhou, Y., Liu, Q. and Tromp, J., 2011, Surface wave sensitivity: mode summation versus adjoint SEM, *Geophys. J. Int.*, **187,** 1560–1576.