How AVO information can be practically incorporated in full waveform inversion

Kris Innanen

ABSTRACT

The claim made in a companion paper, namely that certain formulas for multiparameter reflection full waveform inversion are easy to analyze as well as implement, is robustly challenged in this paper. We ask a question we think is, in fact, a central matter in the future of FWI as applied to pre-critical reflection seismic data (which is the most common and cost-effective kind of seismic data we collect). If FWI automatically converges to the right answer for a particular parameter, through operations on data which are mixtures of the effects of several parameters, it must do some kind of "unmixing" akin to that in AVO inversion. If it cannot, there is no solution to what FWI practitioners refer to as parameter cross-talk. How does this happen? Is it automatic to any FWI procedure—gradient based, quasi-Newton, and Newton alike? Or do we need to properly pose the problem to manage multiple parameters? In this paper we parse our quasi-Newton update formulas, seeking (1) *the internal ability to diagnose ill-posedness*, (2) *the ability to produce balanced updates using different subsets of data*, and (3) *the ability to suppress parameter cross-talk* within them. We ultimately conclude that the quasi-Newton update formula we refer to as the parameter-type approximation is properly equipped to incorporate our basic ideas of AVO inversion into FWI.

INTRODUCTION

In a companion paper (Innanen, 2013), we developed a hybrid continuous-discrete approach to multiparameter reflection full waveform inversion (FWI). In particular, we arrived at a set of three types of approximate Hessian, each leading to a different quasi-Newton update. Of these, the *parameter-type* approximation appears to have a particularly valuable role to play in coping with data caused by variations in multiple parameters. Algorithms with an inability to cope with multiple parameters are said to suffer from parameter *cross-talk*. The purpose of this paper is to evaluate in a quantitative way whether or not, and to what extent, the quasi-Newton update formulas naturally emerging from the hybrid approach successfully suppress parameter cross-talk, and manage related issues.

In a pre-critical reflection experimental configuration, the question *How are multiple parameter variations managed in FWI?* is more or less the same as the question *How could the ideas of AVO/AVA analysis be properly incorporated into FWI?* We are the inheritors of several decades of accumulated AVO wisdom, telling us what information is reliably extractable through seismic amplitude analysis, what is not, and how to go about getting it (Castagna and Backus, 1993; Castagna et al., 1998; Foster et al., 2010). It is very difficult to conceive of a successful reflection mode FWI procedure which does not take maximum advantage of this knowledge base. Thus simply being able to discuss FWI in the language of AVO, and vice versa, would be a contribution, and we set it as one aim of this paper.

To analyze the multiparameter quasi-Newton update formulas, we apply them to an idealized problem: the reconstruction of a 1D Earth model in which the acoustic parameters κ (the bulk modulus) and ρ (the density) vary across a single horizontal interface. There are three reasons for this choice. First, in this configuration there is no phase mismatch between measured and modelled fields, and so it effectively isolates the AVO aspect of the problem of interest to us—i.e., the task of transforming the angle dependent reflection coefficient data into correctly-scaled updates. Second, in a problem this simple, every ingredient, including the data (see also, e.g., Weglein et al., 1986), can be supplied analytically, and in terms of the ideal results, which in this case are the step-lengths taking us to the correct answer. This allows us to create easy-to-analyze relationships between the desired answer (i.e., ideal updates taking us straight to the right answer) and the updates resulting from our hybrid scheme. And third, the acoustic two-parameter problem, being significantly simpler than the three-parameter problem, but containing the same ideas, acts as a manageable proxy for the elastic reflection model which must ultimately be employed.

Through examination of analytic examples, we have identified several important traits in the parameter-type approximation:

1. *Computational tractability.* Suppose in a given problem there are N pixels and M parameters (typically $N > 10^6$ and $M \approx 3$). The full Newton step then involves the inversion of an $NM \times NM$ Hessian matrix for each iterate. In contrast, each iterate of the *parameter-type* quasi-Newton step involves the inversion of an $M \times M$ matrix for each of the N output pixels. This is on the same order of computational complexity as a gradient-based update.

2. *The internal ability to diagnose ill-posedness.* Reflection data corresponding to, for instance, a single fixed slowness or incidence angle, cannot constrain variations in more than one parameter. Neither gradient based steps, nor steps involving the diagonal elements of the Hessian, have internal checks to confirm that the input data are in this sense sufficient. Off-diagonal Hessian elements, of the type retained in the parameter-type quasi-Newton update, supply this check.

3. *The ability to produce stable updates using different subsets of input data.* Convergence of the iterative FWI procedure is hindered if by using, for instance, different groups of slowness or angle values, different step-lengths are determined. The parameter-type quasi-Newton update contains the components of the inverse Hessian which act to balance the update across angles.

4. *The ability to suppress parameter cross-talk.* Seismic data are generally the result of multiple parameters varying in the same volume of the Earth. Determining which parameters must have varied, and how they varied, to produce a given datum is a challenging unmixing problem. Direct seismic inversion research, often based on inverse scattering (Raz, 1981; Clayton and Stolt, 1981; Stolt and Weglein, 1985; Weglein et al., 2003; Innanen and Weglein, 2007) has been primarily concerned with discovering ways multiple parameters can be individually determined from such mixed data. FWI gradients are totally insensitive to this issue; a density gradient, to take a concrete example, will be nonzero even if the density remains constant throughout the Earth, and the data in question are caused only by variations in other parameters.

This is also true of quasi-Newton updates involving only the diagonal elements of the Hessian. This means both gradient-based methods and diagonal quasi-Newton update methods are completely unprotected from parameter cross-talk. In a full waveform inversion iteration, only off-diagonal Hessian elements can in principle combat cross-talk; the parameter-type quasi-Newton update involves these elements.

We will demonstrate that the parameter-type update has these traits, one by one. We will do so by doing analytic calculations of the first iteration of a reflection FWI procedure for low-angle reflection data, assuming a homogeneous background model. We begin by assembling the ingredients we need to do these calculations. We may then quantitatively evaluate the relative effectiveness of quasi-Newton and gradient-based update formulas.

When the analytic data are substituted into the update formulas, the left hand sides and the right hand sides of the resulting equations both in the end contain update quantities, the δs_κ and the δs_ρ to be added to the background models to complete the iteration. The updates on the left hand side are those created by the quasi-Newton formulas. The updates on the right are, as we shall see, the ideal updates taking us directly to the correct answer. Leftover factors in these equations, therefore (e.g., angle dependent coefficients, terms proportional to the wrong update, etc.), moving the relations away from straight statements of equivalence, can be interpreted as flaws in the update formulas. A formula will be "more successful" the less of these factors there are. We will end by summarizing findings in the context of the list above.

ACOUSTIC AVO EQUATIONS IN THE LANGUAGE OF FWI

We will pose the reflection FWI problem in a way particularly suited to incorporating AVO, but, to meet FWI halfway, we will also express our familiar AVO/AVA equations in convenient, if slightly different FWI terms. This will easily allow them to be used to analyze the FWI step.

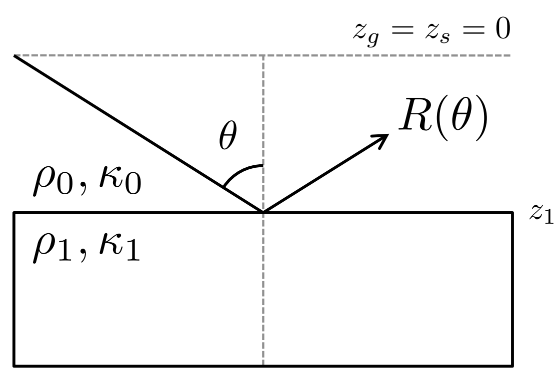


FIG. 1. Geometry of a reflection coming from a single horizontal planar acoustic interface.

The acoustic reflection coefficient associated with a plane wave impinging at angle θ on a horizontal interface which separates an upper medium with density and bulk modulus values ρ_0 and κ_0 from a lower medium with values ρ_1 and κ_1 (Figure 1) is given by

$$R(\theta) = \frac{1 - \Omega(\theta)}{1 + \Omega(\theta)}, \tag{1}$$

where

$$\Omega(\theta) = \left(\frac{\rho_0}{\rho_1}\right)\left(\frac{\kappa_0}{\kappa_1}\frac{\rho_1}{\rho_0}\right)^{1/2}\left(1 - \sin^2\theta\right)^{-1/2}\left(1 - \frac{\kappa_1}{\kappa_0}\frac{\rho_0}{\rho_1}\sin^2\theta\right)^{1/2}. \tag{2}$$

Defining the reciprocal modulus and density parameters

$$s_\kappa = \frac{1}{\kappa_1}, \quad s_{\kappa_0} = \frac{1}{\kappa_0}, \quad s_\rho = \frac{1}{\rho_1}, \quad s_{\rho_0} = \frac{1}{\rho_0}, \tag{3}$$

the reflection coefficient, to first order in the changes

$$\delta s_\kappa = s_\kappa - s_{\kappa_0}, \quad \delta s_\rho = s_\rho - s_{\rho_0}, \tag{4}$$

and $\sin^2\theta$, is

$$R(\theta) \approx -\frac{1}{4}\frac{1}{\cos^2\theta}\left(\frac{\delta s_\kappa}{s_{\kappa_0}}\right) - \frac{1}{4}\cos^2\theta\left(\frac{\delta s_\rho}{s_{\rho_0}}\right). \tag{5}$$

The significance of this to our understanding of reflection full waveform inversion is as follows. If the initial, or background medium is homogeneous, and the data are due to a single acoustic interface at pre-critical angles, then relation (5) holds for *ideal, perfect* first updates δs_κ and δs_ρ. When we later devise formulas for the gradients and Hessian functions, and consider data of this type as input, we can substitute the form in (5) and see how close to these ideal updates we have come.

ANALYSIS OF THE RECONSTRUCTION OF AN ACOUSTIC BOUNDARY

The formulas in the companion paper (Innanen, 2013) have going for them that they expose to the eye parts of the machinery of a Newton step in full waveform inversion. A characterization of how this machinery copes with multiparameter reflection issues is most clearly accomplished by working a specific example. We will consider the problem of reconstruction of a single acoustic boundary. We will require the appropriate equations of motion, spaces in which to analyze the data, wave solutions in the chosen background medium, and analytic data sets to proceed. Let us assemble these one by one.

Ingredients

Acoustic equations

We will assume that our data are measurements of an acoustic field which satisfies

$$\left[\nabla \cdot \left(\frac{1}{\rho(\mathbf{r})}\right)\nabla + \frac{\omega^2}{\kappa(\mathbf{r})}\right] P(\mathbf{r}, \mathbf{r}_s, \omega) = \delta(\mathbf{r} - \mathbf{r}_s), \tag{6}$$

where $\mathbf{r}$ is the observation position vector in 1D, 2D, or 3D as needed, $\mathbf{r}_s$ is the location of an idealized impulsive source, and ω is the angular frequency. The medium is characterized by spatial variations in the mass density ρ and the bulk modulus κ. For convenience we deal not with these parameters explicitly but with their reciprocals s_κ and s_ρ:

$$\left[\nabla \cdot s_\rho(\mathbf{r})\nabla + \omega^2 s_\kappa(\mathbf{r})\right] P(\mathbf{r}, \mathbf{r}_s, \omega) = \delta(\mathbf{r} - \mathbf{r}_s). \tag{7}$$

At any given step in a full waveform inversion procedure we work towards updating a model (s_κ, s_ρ) by and amount $(\delta s_\kappa, \delta s_\rho)$ and adding it to a current iterate $(s_{\kappa 0}, s_{\rho 0})$. That is,

$$\begin{aligned} s_\rho(\mathbf{r}) &= s_{\rho 0}(\mathbf{r}) + \delta s_\rho(\mathbf{r}), \\ s_\kappa(\mathbf{r}) &= s_{\kappa 0}(\mathbf{r}) + \delta s_\kappa(\mathbf{r}). \end{aligned} \tag{8}$$

The current model parameters $s_{\kappa 0}$ and $s_{\rho 0}$ are knowns, as are the solutions for the fields in that medium, namely G where

$$\left[\nabla \cdot s_{\rho 0}(\mathbf{r})\nabla + \omega^2 s_{\kappa 0}(\mathbf{r})\right] G(\mathbf{r}, \mathbf{r}_s, \omega) = \delta(\mathbf{r} - \mathbf{r}_s). \tag{9}$$

The updates $(\delta s_\kappa, \delta s_\rho)$ are the unknowns to be determined by .

The (k_g, x_s, ω) domain

We will work the example in 1.5D, i.e., assuming a medium which varies in depth z, overlain by line sources and receivers arranged along a lateral x axis. To facilitate the eventual analysis of the update in AVO/AVA terms, i.e., such that it involves incidence angles etc., it is convenient to begin by considering the data and modeled wave fields in the (k_g, x_s, ω) domain, where k_g is the Fourier conjugate to the lateral receiver coordinate x_g and x_s is the lateral source coordinate. The wave field P observed at depth z, due to a source at (x_s, z_s), in these coordinates will therefore be expressed as

$$P(k_g, z, x_s, z_s, \omega). \tag{10}$$

During the nth FWI iteration we assume access to the modelled field G everywhere in the nth medium iterate. At depth z we will write G as

$$G(k_g, z, x_s, z_s, \omega | s_\kappa^{(n)}, s_\rho^{(n)}), \tag{11}$$

where we explicitly refer to the dependence of the modelled field on the medium properties.

Green's functions for a homogeneous initial medium

We will now restrict ourselves to the first iteration of the FWI problem, and assume a homogeneous acoustic medium as the background. This means we have access to analytic, closed form solutions for the modelled field G. Solving equation (9) with spatially constant $s_{\kappa 0}$ and $s_{\rho 0}$ (e.g., Clayton and Stolt, 1981) we have two useful forms

$$G_0(k_g, z, x_s, z_s, \omega) = \rho_0 e^{-ik_g x_s} \frac{e^{iq_g|z-z_s|}}{i2q_g}, \tag{12}$$

and

$$G_0(x_g, z, x_s, z_s, \omega) = \frac{\rho_0}{2\pi} \int_{-\infty}^{\infty} dk' e^{ik'(x_g - x_s)} \frac{e^{iq'|z-z_s|}}{i2q'}, \tag{13}$$

where

$$q_g = \frac{\omega}{c_0}\sqrt{1 - \frac{k_g^2 c_0^2}{\omega^2}}, \quad q' = \frac{\omega}{c_0}\sqrt{1 - \frac{k'^2 c_0^2}{\omega^2}}. \tag{14}$$

Here $c_0 = \sqrt{\kappa_0/\rho_0} = \sqrt{s_{\rho 0}/s_{\kappa 0}}$ is the constant background acoustic wave velocity.

Analytic data and residuals

The data set, $D = P(k_g, z_g, x_s, z_s, \omega)$, i.e., the wave field evaluated on the surface $z = z_g$, at its most general is expressed as

$$D(k_g, z_g, x_s, z_s, \omega). \tag{15}$$

Let us now further assume the observed data are due to a reflection from a single horizontal interface, which can be written analytically. Since the modelled field at the first iteration can also be written analytically, so too can the $n = 0$ residuals, i.e., the difference between the two. Without loss of generality we can place the source at the origin, $z_s = x_s = 0$, and, letting that origin lie on the measurement surface, let $z_g = 0$ also. The data measured above the single interface at depth z_1 are given by

$$D(k_g, 0, 0, 0, \omega) = \frac{1}{i2q_g} + R(\theta)\frac{e^{i2q_g z_1}}{i2q_g}, \tag{16}$$

where the first term is the "direct" component of the wave field, propagating from the source to the receiver (which are coincident in depth in this example), and the second term is the reflection coming from the interface at depth z_1. The form of the reflection coefficient $R(\theta)$ was given exactly in equation (1) and approximately in equation (2).

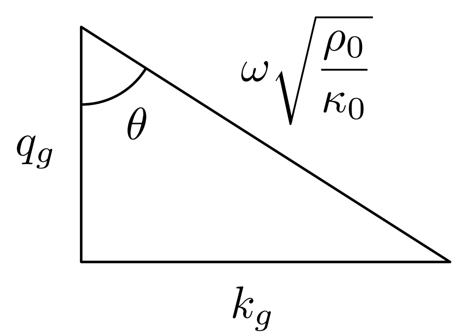


FIG. 2. Geometry of harmonic plane wavenumbers, frequencies, and angles.

In Figure 2 we recall the harmonic plane wave relationships by which the angle in the reflection coefficient is related to the Fourier quantities k_g, q_g and ω in the data and wave field solution expressions. This geometry provides the following useful relations:

$$\tan\theta = \frac{k_g}{q_g}, \quad \cos\theta = \frac{q_g}{\omega\sqrt{\rho_0/\kappa_0}} = \frac{q_g c_0}{\omega}, \tag{17}$$

and

$$d\omega = d(-2q_g)\left(-\frac{c_0}{2\cos\theta}\right). \tag{18}$$

Let us finally express the residuals associated with the first FWI iterate. The modelled field is the same as the first term in the data in equation (16), the direct wave being the full solution in a homogeneous medium:

$$G(k_g, 0, 0, 0, \omega | s_\kappa^{(0)}, s_\rho^{(0)}) = \frac{1}{i2q_g}. \tag{19}$$

The calculation of the gradient and Hessian functions requires the complex conjugate of the residuals, which is the complex conjugate of equation (19) subtracted from the complex conjugate of equation (16):

$$\delta P^*(k_g, \omega | s_\kappa^{(0)}, s_\rho^{(0)}) = -R(\theta)\frac{e^{-i2q_g z_1}}{i2q_g}. \tag{20}$$

Acoustic gradients for depth-varying models

A (k_g, ω) domain objective function

The gradient and Hessian functions we will construct characterize the step direction towards a local minimum of the objective function. The objective function itself is designed to penalize data misfit, and, often, models of an undesired type. Here we will use a simple data misfit objective function. To match with the naming convention used by Innanen (2013), in this two parameter case we refer to the objective function as Φ_2, where

$$\Phi_2\left(s_\kappa, s_\rho\right) = \frac{1}{2}\sum_{k_g}\int d\omega |\delta P|^2, \tag{21}$$

which measures the squared difference between measured and modelled data, summed over the experimental variables k_g and ω. In keeping with the hybrid discrete-continuous formulation we let the sum be continuous over ω but discrete over k_g.

Bulk modulus and density gradients for depth-varying models

We next restrict the general FWI update problem further, permitting the unknowns to be depth varying only. Standard analysis of objective functions of the form in equation (21) leads to forms for the gradients, namely the derivatives of Φ_2 with respect to each of the two acoustic parameters $s_\kappa(z)$ and $s_\rho(z)$. We have

$$g_\kappa^{(n)}(z) = -\sum_{k_g}\int d\omega \frac{\partial G(k_g, z_g, x_s, z_s, \omega | s_\kappa^{(n)}, s_\rho^{(n)})}{\partial s_\kappa^{(n)}(z)} \delta P^*(k_g, z_g, x_s, z_s, \omega | s_\kappa^{(n)}, s_\rho^{(n)}), \tag{22}$$

and

$$g_\rho^{(n)}(z) = -\sum_{k_g}\int d\omega \frac{\partial G(k_g, z_g, x_s, z_s, \omega | s_\kappa^{(n)}, s_\rho^{(n)})}{\partial s_\rho^{(n)}(z)} \delta P^*(k_g, z_g, x_s, z_s, \omega | s_\kappa^{(n)}, s_\rho^{(n)}). \tag{23}$$

The gradients, in other words, are constructed from two main quantities, the sensitivities $\partial G/\partial s$ and the residuals δP. Perturbation theory or an adjoint state approach may additionally be applied to find forms for the sensitivities. The gradient for updating in the modulus parameter is

$$g_\kappa^{(n)}(z) = \sum_{k_g} \int d\omega \omega^2 \int dx' G(k_g, z_g, x', z, \omega | s_\kappa^{(n)}, s_\rho^{(n)}) G(x', z, x_s, z_s, \omega | s_\kappa^{(n)}, s_\rho^{(n)}) \delta P^*,$$

and the gradient for updating the density parameter is

$$g_\rho^{(n)}(z) = -\sum_{k_g} \int d\omega \int dx' \left[\frac{\partial G(k_g, z_g, x', z, \omega | s_\kappa^{(n)}, s_\rho^{(n)})}{\partial z} \frac{\partial G(x', z, x_s, z_s, \omega | s_\kappa^{(n)}, s_\rho^{(n)})}{\partial z} \right.$$
$$\left. - G(k_g, z_g, x', z, \omega | s_\kappa^{(n)}, s_\rho^{(n)}) \frac{\partial^2 G(x', z, x_s, z_s, \omega | s_\kappa^{(n)}, s_\rho^{(n)})}{\partial x'^2} \right] \delta P^*.$$

In the first iteration the modelled fields G have analytic forms. By substituting the Green's functions in equations (12) and (14) into these formulas we obtain for the $n = 0$ case

$$g_\kappa^{(0)}(z) = -\frac{1}{4} \sum_{k_g} \int d\omega \; e^{-iq_g(z_g+z_s)-ik_g x_s} e^{i2q_g z} \left(\frac{\omega^2}{q_g^2} \right) \delta P^*, \quad (24)$$

for the modulus update, and

$$g_\rho^{(0)}(z) = -\frac{1}{4} \sum_{k_g} \int d\omega \; e^{-iq_g(z_g+z_s)-ik_g x_s} e^{i2q_g z} \left(1 - \frac{k_g^2}{q_g^2} \right) \delta P^*, \quad (25)$$

for the density update.

Transformation from k_g to angle

In our development of the single-interface case we have already set $z_g = z_s = x_s = 0$, which will simplify equations (24)-(25). We may additionally now easily transform the sum over k_g to a sum over incidence angle, using the relations in equations (17):

$$g_\kappa^{(0)}(z) = -\frac{c_0^2}{4} \sum_\theta \int d\omega \; e^{i2q_g z} \left(\frac{1}{\cos^2\theta} \right) \delta P^*, \quad (26)$$

and

$$g_\rho^{(0)}(z) = -\frac{1}{4} \sum_\theta \int d\omega \; e^{i2q_g z} \left(1 - \tan^2\theta \right) \delta P^*, \quad (27)$$

for the modulus and density gradients respectively.

Two criticisms of gradient-based methods in multiparameter reflection FWI

With the gradients in equations (26)-(27) we can carry out a valid form of FWI of the *gradient-based* kind. With as little as one angle of data, say θ_1, the steps

$$\delta s_\kappa(z) = \mu_\kappa g_\kappa^{(0)}(z|\theta_1) = -\frac{c_0^2}{4}\int d\omega \; e^{i2q_g z}\left(\frac{1}{\cos^2\theta_1}\right)\delta P^* \tag{28}$$

and

$$\delta s_\rho(z) = \mu_\rho g_\rho^{(0)}(z|\theta_1) = -\frac{1}{4}\int d\omega \; e^{i2q_g z}\left(1-\tan^2\theta_1\right)\delta P^* \tag{29}$$

where the scalars μ_κ and μ_ρ are calculated via a line search can be added to the current model iterates to complete an update. For the special case of a single acoustic interface, by substituting the analytic data residuals in equation (16) into the modulus gradient formula we obtain

$$\begin{aligned} g_\kappa^{(0)}(z|\theta_1) &= -\frac{c_0^2}{4}\int d\omega \; e^{i2q_g z}\left(\frac{1}{\cos^2\theta_1}\right)\left(-R(\theta_1)\frac{e^{-i2q_g z_1}}{i2q_g}\right) \\ &= \frac{R(\theta_1)c_0^2}{4\cos^2\theta_1}\int d\omega \;\left(\frac{e^{i2q_g(z-z_1)}}{i2q_g}\right) \\ &= -\frac{R(\theta_1)c_0^3}{8\cos^3\theta_1}\int d(-2q_g)\;\left(\frac{e^{i2q_g(z-z_1)}}{i2q_g}\right) \\ &= -\frac{R(\theta_1)c_0^3}{8\cos^3\theta_1}S(z-z_1), \end{aligned} \tag{30}$$

where in the second last step we have used the plane-wave geometrical relationship in equation (18), and in the last step we have recognized the integral as the inverse Fourier transform of the spectrum of the step function S where

$$S(z) = \begin{cases} 0, & z < 0 \\ 1, & z > 0 \end{cases}, \tag{31}$$

with its step location at z_1. Similarly the density gradient is found to be

$$\begin{aligned} g_\rho^{(0)}(z|\theta_1) &= -\frac{R(\theta_1)c_0(1-\tan^2\theta_1)}{8\cos\theta_1}S(z-z_1) \\ &\approx -\frac{R(\theta_1)c_0\cos\theta_1}{8}S(z-z_1), \end{aligned} \tag{32}$$

the last approximation being appropriate for small angles in which $\tan\theta \approx \sin\theta$. It is fair to expect *a priori* that a two-parameter inverse problem cannot be solved with a single angle of data, though we have not specifically demonstrated that this is the case in our current formulation. Nevertheless, to accommodate the possibility that N angles of data are needed, we can extend the formula for the set Θ where

$$\Theta = \{\theta_1, \theta_2, ..., \theta_N\}, \tag{33}$$

as follows:

$$
\begin{aligned}
g_{\kappa}^{(0)}(z|\Theta) &= -\frac{c_0^3}{8}\left[\sum_j \frac{R(\theta_j)}{\cos^3\theta_j}\right] S(z-z_1) \\
g_{\rho}^{(0)}(z|\Theta) &= -\frac{c_0}{8}\left[\sum_j R(\theta_j)\cos\theta_j\right] S(z-z_1).
\end{aligned}
\tag{34}
$$

We have the general mathematical expectation that the minimum number of angles needed to determine two parameter contrasts is two. Assuming this holds for FWI updates, the simplest well-posed single interface update must involve a two-angle objective function and the resulting two-angle gradients

$$
\begin{aligned}
g_{\kappa}^{(0)}(z|\theta_1,\theta_2) &= -\frac{c_0^3}{8}\left[\frac{R(\theta_1)}{\cos^3\theta_1}+\frac{R(\theta_2)}{\cos^3\theta_2}\right] S(z-z_1) \\
g_{\rho}^{(0)}(z|\theta_1,\theta_2) &= -\frac{c_0}{8}\left[R(\theta_1)\cos\theta_1+R(\theta_2)\cos\theta_2\right] S(z-z_1).
\end{aligned}
\tag{35}
$$

How "good" is a proposed update?

Let us devise a way of evaluating the "goodness" of a proposed gradient-based, quasi-Newton, or Newton update, using our analytic results. The gradients in equations (35) are functions of the reflection coefficient $R(\theta)$; but the reflection coefficient itself can be related to the idealized updates, as we developed in equation (5):

$$
R(\theta) \approx -\frac{1}{4s_{\kappa_0}}\left(\frac{1}{\cos^2\theta}\right)\delta s_{\kappa} - \frac{1}{4s_{\rho_0}}\left(\cos^2\theta\right)\delta s_{\rho}. \tag{36}
$$

Substituting equation (36) into proposed update formulas like those in equation (35) produces expressions relating the proposed updates (on the left hand sides) to the idealized updates (on the right hand sides). The "better" a proposed update is, the closer this relationship will be to a straight equality: proposed update = idealized update. Because equation (36) is a first order approximation of R, the evaluations we make are also correct to first order.

To be specific, the best possible result to find for an update formula under study would be that the right hand side of, for instance, the κ formula in equation (35) was found to be

$$
\delta s_{\kappa} S(z-z_1), \tag{37}
$$

i.e., with the step at the correct depth z_1 and with an amplitude exactly equal to the ideal length as found in equation (36). A close second would be to find that the proposed update was within a scalar multiple of equation (37). A distant third would be to find that the proposed update was equal to equation (37) but only to within a coefficient which depends on an experimental variable like θ. And a very distant last place finish would be to find that the proposed κ update was proportional to δs_{κ} *and* δs_{ρ}, which would indicate significant cross-talk. Our scheme in order of best to worst:

1. The update is found to be $\delta s_\kappa S(z - z_1)$.
2. The update is within a scalar multiple of $\delta s_\kappa S(z - z_1)$.
3. The update is explicitly dependent on experimental parameters, e.g., θ.
4. The κ update formula is a function of both δs_κ and δs_ρ.

Criticism I: angle-dependent updates

On this evaluation scale, gradient-based methods achieve a (3.) or a (4.), and the third place finish only arises in highly contrived special cases. Let us begin with these easy special cases, in which the reflecting interface is known to involve a contrast in just one of the two parameters. If the contrast is in bulk modulus, R is then approximately simplified to

$$R(\theta) \approx -\frac{1}{4s_{\kappa 0}}\left(\frac{1}{\cos^2\theta}\right)\delta s_\kappa. \tag{38}$$

Whereas if the reflection is due to a contrast in density only, R is

$$R(\theta) \approx -\frac{1}{4s_{\rho 0}}\left(\cos^2\theta\right)\delta s_\rho. \tag{39}$$

Substituting equation (38) into the κ gradient formula in equation (35), we have

$$g_\kappa^{(0)}(z|\theta_1,\theta_2) \approx \frac{c_0^3}{32}\left[\frac{1}{\cos^5\theta_1} + \frac{1}{\cos^5\theta_2}\right]\left(\frac{\delta s_\kappa}{s_{\kappa 0}}\right) S(z - z_1). \tag{40}$$

Likewise if the data involve the reflection coefficient in equation (39), the density gradient formula simplifies to

$$g_\rho^{(0)}(z|\theta_1,\theta_2) \approx \frac{c_0}{32}\left[\cos^3\theta_1 + \cos^3\theta_2\right]\left(\frac{\delta s_\rho}{s_{\rho 0}}\right) S(z - z_1). \tag{41}$$

How close have we come to the right answer? Comparing, for instance, equation (40) to equation (37), we see that the correct depth of the step is produced. This is because the initial FWI medium and the actual medium between the source/receiver plane and the interface are in agreement. The major difference is that the update formula is dependent on the two angles of data we have used. This is a significant flaw, as the actual update has no such dependence. A more useful formula would contain pre-factors which were themselves functions of the input angles, and whose effect, when in a product with data amplitudes, would be to suppress the overall angle dependence of the update.

Criticism II: bulk modulus/density cross-talk

Let us consider a much more significant issue — namely what is the bulk modulus gradient when the data are due to a contrast in density only? The correct answer is zero, but inspection of equations (35) gives very little confidence that this will be true. Let us make this second criticism in more general terms. If both parameters change across the interface, which is by far the most likely scenario, the gradients in equation (35) become

$$\begin{aligned} g_{\kappa}^{(0)}(z|\theta_1,\theta_2) \approx & \frac{c_0^3}{32}\left[\frac{1}{\cos^5\theta_1}+\frac{1}{\cos^5\theta_2}\right]\left(\frac{\delta s_\kappa}{s_{\kappa 0}}\right)S(z-z_1) \\ & + \frac{c_0^3}{32}\left[\frac{1}{\cos\theta_1}+\frac{1}{\cos\theta_2}\right]\left(\frac{\delta s_\rho}{s_{\rho 0}}\right)S(z-z_1), \end{aligned} \tag{42}$$

and

$$\begin{aligned} g_{\rho}^{(0)}(z|\theta_1,\theta_2) \approx & \frac{c_0}{32}\left[\cos^3\theta_1+\cos^3\theta_2\right]\left(\frac{\delta s_\rho}{s_{\rho 0}}\right)S(z-z_1) \\ & + \frac{c_0}{32}\left[\frac{1}{\cos\theta_1}+\frac{1}{\cos\theta_2}\right]\left(\frac{\delta s_\kappa}{s_{\kappa 0}}\right)S(z-z_1), \end{aligned} \tag{43}$$

in other words both updates are proportional to *both* ideal updates: a distant fourth place finish. The presence of significant parameter cross-talk—in which there is algorithmic uncertainty about to which contrasts caused which data variations—is indicated.

Acoustic Hessian functions for depth-varying models

Qualitative inspection of the quasi-Newton update formulas hint that the problems plaguing gradient based methods might be mitigated by the involvement of the Hessian. We will now test the extent to which this is true. We start by computing the four two-parameter Hessian functions as described in the first of these two papers (Innanen, 2013). Defining for convenience

$$\mathrm{T}=e^{iq_g(z+|z-z'|+z')},\quad \mathrm{U}=e^{-iq_g(z_g+z_s)-ik_gx_s},\quad \mathrm{Z}=\mathrm{sgn}(z-z'), \tag{44}$$

we have the four functions $H_{\kappa\kappa}^{(n)}(z,z')$, $H_{\kappa\rho}^{(n)}(z,z')$, $H_{\rho\kappa}^{(n)}(z,z')$ and $H_{\rho\rho}^{(n)}(z,z')$ to determine for the $n=0$ case, given the details of this single acoustic interface problem with its associated analytic data. The κ-κ function at iteration n is

$$H_{\kappa\kappa}^{(n)}(z,z')=\frac{\partial^2\Phi_2(s_\kappa^{(n)},s_\rho^{(n)})}{\partial s_\kappa^{(n)}(z)\partial s_\kappa^{(n)}(z')}=\frac{\partial}{\partial s_\kappa^{(n)}(z)}g_\kappa^{(n)}(z'). \tag{45}$$

Using the form for the gradient determined in the previous section, and again invoking either a perturbation theoretic framework or adjoint-state methods, after some manipulation we find for $n=0$

$$H_{\kappa\kappa}^{(0)}(z,z')=2\sum_{k_g}\int d\omega\frac{\omega^4}{(i2q_g)^3}\left[\mathrm{U}\mathrm{T}\delta P^*-\frac{e^{i2q_g(z'-z)}}{i2q_g}\right].$$

Similarly for the κ-ρ and ρ-κ cases we have

$$H_{\kappa\rho}^{(n)}(z,z')=\frac{\partial^2\Phi_2(s_\kappa^{(n)},s_\rho^{(n)})}{\partial s_\kappa^{(n)}(z)\partial s_\rho^{(n)}(z')}=\frac{\partial}{\partial s_\kappa^{(n)}(z)}g_\rho^{(n)}(z'), \tag{46}$$

which for the $n = 0$ case leads to

$$H^{(0)}_{\kappa\rho}(z,z') = 2\sum_{k_g}\int d\omega \frac{\omega^2}{(i2q_g)^3}\left[\mathrm{UT}\left(\mathrm{Z}q_g^2 - k_g^2\right)\delta P^* + \frac{1}{2}\frac{e^{i2q_g(z'-z)}}{i2q_g}(q_g^2 - k_g^2)\right],$$

and

$$H^{(n)}_{\rho\kappa}(z,z') = \frac{\partial^2\Phi_2(s^{(n)}_\kappa, s^{(n)}_\rho)}{\partial s^{(n)}_\rho(z)\partial s^{(n)}_\kappa(z')} = \frac{\partial}{\partial s^{(n)}_\rho(z)} g^{(n)}_\kappa(z'). \tag{47}$$

such that at $n = 0$ we obtain

$$H^{(0)}_{\rho\kappa}(z,z') = 2\sum_{k_g}\int d\omega \frac{\omega^2}{(i2q_g)^3}\left[\mathrm{UT}\left(\mathrm{Z}q_g^2 - k_g^2\right)\delta P^* + \frac{1}{2}\frac{e^{i2q_g(z'-z)}}{i2q_g}(q_g^2 - k_g^2)\right],$$

confirming the general expectation that $H^{(0)}_{\kappa\rho} = H^{(0)}_{\rho\kappa}$. Finally for the ρ-ρ case, beginning with the general form

$$H^{(n)}_{\rho\rho}(z,z') = \frac{\partial^2\Phi_2(s^{(n)}_\kappa, s^{(n)}_\rho)}{\partial s^{(n)}_\rho(z)\partial s^{(n)}_\rho(z')} = \frac{\partial}{\partial s^{(n)}_\rho(z)} g^{(n)}_\rho(z'), \tag{48}$$

we obtain at iteration $n = 0$

$$H^{(0)}_{\rho\rho}(z,z') = -2\sum_{k_g}\int \frac{d\omega}{(i2q_g)^3}\left[\mathrm{UT}\left(k_g^4 - \mathrm{Z}2k_g^2q_g^2 - q_g^4\right)\delta P^* + \frac{1}{2}\frac{e^{i2q_g(z'-z)}}{i2q_g}(k_g^2 - q_g^2)^2\right].$$

These explicit expressions may now be used in our analysis of the Newton and quasi-Newton update formulas.

Transformation from k_g to angle

As in the gradient calcualtion, we are attempting to understand an FWI update while simultaneously phrasing the results in a form familiar to practitioners and theoreticians of seismic AVO and AVA. So, at this stage, we again use the relations in equations (17)-(18), to re-write the Hessian functions as

$$H^{(0)}_{\kappa\kappa}(z,z') = \sum_\theta \frac{c_0^4}{8\cos^4\theta}\int d\omega\left[e^{i2q_g(z'-z)} - (i2q_g)\mathrm{T}\delta P^*\right], \tag{49}$$

and

$$H^{(0)}_{\kappa\rho}(z,z') = H^{(0)}_{\rho\kappa}(z,z') \approx \frac{c_0^2}{16}\sum_\theta\int d\omega\left[e^{i2q_g(z-z')} + \mathrm{ZT}(i2q_g)\frac{[1+\mathrm{Z}\tan^2\theta]}{\cos^2\theta}\delta P^*\right], \tag{50}$$

and

$$H^{(0)}_{\rho\rho}(z,z') \approx \sum_\theta \frac{\cos^4\theta}{16}\int d\omega\left[e^{i2q_g(z'-z)} + \mathrm{T}(i2q_g)\frac{[1-2\mathrm{Z}\tan^2\theta - \tan^4\theta]}{\cos^4\theta}\delta P^*\right]. \tag{51}$$

Notice that the Hessian functions have been naturally decomposed into residual-independent parts (without the term δP^*) and the residual-dependent parts (with the term δP^*).

Parameter-type and combined-type approximate Hessian functions

If we wish to follow up on the provisional claim that a *parameter-type* Hessian approximation suppresses cross-talk, our remaining task is as follows. The approximation relies on the idea that the Hessian functions can be written

$$\begin{aligned} H_{\kappa\kappa}(z,z') &\approx \Gamma_{\kappa\kappa}(z)\delta(z-z'), \quad H_{\kappa\rho}(z,z') \approx \Gamma_{\kappa\rho}(z)\delta(z-z'), \\ H_{\rho\kappa}(z,z') &\approx \Gamma_{\rho\kappa}(z)\delta(z-z'), \quad H_{\rho\rho}(z,z') \approx \Gamma_{\rho\rho}(z)\delta(z-z'), \end{aligned} \tag{52}$$

i.e., which are diagonal in their space-dependence—i.e., zero unless $z = z'$. We must therefore identify parts of the Hessian functions in equations (49)–(51) whose space dependences are delta functions, whereupon the coefficients $\Gamma_{\kappa\kappa}$, $\Gamma_{\kappa\rho}$, $\Gamma_{\rho\kappa}$ and $\Gamma_{\rho\rho}$ can be found. It is not difficult to show that the residual-independent parts of the Hessian functions fit this bill. For instance, in the κ-κ case, we find by neglecting the term proportional to δP^*,

$$\begin{aligned} H_{\kappa\kappa}^{(0)}(z,z') &= \sum_\theta \frac{c_0^4}{8\cos^4\theta}\int d\omega \left[e^{i2q_g(z'-z)} - (i2q_g)\mathrm{T}(z,z')\delta P^*\right] \\ &\approx \sum_\theta \frac{c_0^4}{8\cos^4\theta}\int d\omega e^{i2q_g(z'-z)} \\ &\approx -\sum_\theta \frac{c_0^5}{16\cos^5\theta}\int d(-2q_g) e^{i2q_g(z'-z)} \\ &\approx -\sum_\theta \frac{c_0^5}{16\cos^5\theta}\delta(z-z'). \end{aligned} \tag{53}$$

The coefficient $\Gamma_{\kappa\kappa}(z)$ is in the single interface case is simply the constant $\Gamma_{\kappa\kappa}$ where

$$\Gamma_{\kappa\kappa} = -\sum_\theta \frac{c_0^5}{16\cos^5\theta}. \tag{54}$$

Likewise we may find

$$\Gamma_{\kappa\rho} = \Gamma_{\rho\kappa} = -\sum_\theta \frac{c_0^3}{16\cos\theta}, \tag{55}$$

and

$$\Gamma_{\rho\rho}(z) = -\sum_\theta \frac{c_0\cos^3\theta}{16}. \tag{56}$$

With these coefficients and analytic forms for the gradient functions, we have all the necessary ingredients to calculate the *parameter-type* quasi-Newton update as it would arise for data from a single interface:

$$\begin{bmatrix} \delta s_\kappa(z) \\ \delta s_\rho(z) \end{bmatrix} \approx -\frac{1}{\Gamma_{\kappa\kappa}\Gamma_{\rho\rho} - \Gamma_{\rho\kappa}\Gamma_{\kappa\rho}} \begin{bmatrix} \Gamma_{\rho\rho} & -\Gamma_{\rho\kappa} \\ -\Gamma_{\kappa\rho} & \Gamma_{\kappa\kappa} \end{bmatrix} \begin{bmatrix} g_\kappa(z) \\ g_\rho(z) \end{bmatrix}. \tag{57}$$

We can also compute the *combined-type* quasi-Newton update

$$\begin{bmatrix} \delta s_\kappa(z) \\ \delta s_\rho(z) \end{bmatrix} \approx \begin{bmatrix} \Gamma_{\kappa\kappa}^{-1} g_\kappa(z) \\ \Gamma_{\rho\rho}^{-1} g_\rho(z) \end{bmatrix}. \tag{58}$$

Correction for angle-dependent updates with the combined-type approximate Hessian

Let us put the two quasi-Newton update formulas in equations (57)–(58) through the same two tests applied to the gradient based update, the easy test in which one parameter is known to be invariant, and the more difficult test in which both parameters vary simultaneously, bringing about the danger of cross-talk.

First the easy test. If only one parameter varies, and we know that that is the case, the *combined-type* approximate Hessian represented by equation (58) acts to correct the issue of the dependence of the update on experimental variables (in this case, the angle). Again assuming the presence of two angles of data, the coefficient $\Gamma_{\kappa\kappa}$ becomes

$$\Gamma_{\kappa\kappa}(\theta_1,\theta_2) = -\frac{c_0^5}{16}\left[\frac{1}{\cos^5\theta_1}+\frac{1}{\cos^5\theta_2}\right],$$

and the coefficient $\Gamma_{\rho\rho}$ becomes

$$\Gamma_{\rho\rho}(\theta_1,\theta_2) = -\frac{c_0}{16}\left[\cos^3\theta_1+\cos^3\theta_2\right].$$

Now, if indeed $R(\theta)$ is due to a contrast in the modulus only, the gradient will have the form it took in equation (40), namely

$$g_\kappa^{(0)}(z|\theta_1,\theta_2) \approx \frac{c_0^3}{32}\left[\frac{1}{\cos^5\theta_1}+\frac{1}{\cos^5\theta_2}\right]\left(\frac{\delta s_\kappa}{s_{\kappa 0}}\right)S(z-z_1). \quad (59)$$

And, if indeed $R(\theta)$ is due to a contrast in density only, the gradient will be

$$g_\rho^{(0)}(z|\theta_1,\theta_2) \approx \frac{c_0}{32}\left[\cos^3\theta_1+\cos^3\theta_2\right]\left(\frac{\delta s_\rho}{s_{\rho 0}}\right)S(z-z_1). \quad (60)$$

The combined-type quasi-Newton steps for each of these two separate cases are then

$$\Gamma_{\kappa\kappa}^{-1} g_\kappa^{(0)}(z|\theta_1,\theta_2) \approx -\frac{1}{2c_0^2}\left(\frac{\delta s_\rho}{s_{\rho 0}}\right)S(z-z_1), \quad (61)$$

when κ alone varies, and

$$\Gamma_{\rho\rho}^{-1} g_\rho^{(0)}(z|\theta_1,\theta_2) \approx -\frac{1}{2}\left(\frac{\delta s_\rho}{s_{\rho 0}}\right)S(z-z_1) \quad (62)$$

when ρ alone varies. The angle-dependence of the steps in both cases has been suppressed, and, we see that, to first order in δs_κ, δs_ρ and $\sin^2\theta$, the updates are correct regardless of the angle used.

The combined approximation, in other words, achieves a second place finish—but only for the easy of the two tests. The cross-talk test cannot be passed by this approximation: dividing the gradients in equations (42)-(43) through by the factors $\Gamma_{\kappa\kappa}(\theta_1,\theta_2)$ and $\Gamma_{\rho\rho}(\theta_1,\theta_2)$ has no effect on their dependence on both δs_κ and δs_ρ. The combined approximation drops to the bottom of the scale for general multiparameter reflection data, falling prey to cross talk error.

The parameter-type approximate Hessian and ill-posedness

Now let us examine the parameter-type approximate Hessian. An immediate positive consequence of doing so is as follows. Suppose because of lack of data coverage we faced an ill-posed inverse problem. For instance, suppose we were attempting to solve for updates in both modulus and density from a single angle of data, a problem which in standard AVO practice is well known to be underdetermined. If we were to apply the gradient-based method, or indeed use a combined-approximate Hessian, not only would we suffer from cross-talk, but we would receive no indication from the approach that the problem was ill-posed. Poorly-constrained models would be blithely generated by those formulas as if nothing was wrong.

However, if we attempt to solve for *parameter-type* quasi-Newton updates, we are given immediate notice that the problem is ill posed. To see this, suppose we were to try to solve for the density update with the parameter-type formula using one angle of data. The coefficients $\Gamma_{\kappa\kappa}$, $\Gamma_{\kappa\rho}$, $\Gamma_{\rho\kappa}$ and $\Gamma_{\rho\rho}$ given a single angle θ of data, calculated using equations (54)–(56) form the determinant

$$\Gamma_{\kappa\kappa}\Gamma_{\rho\rho} - \Gamma_{\kappa\rho}\Gamma_{\rho\kappa} = \left(-\frac{c_0^5}{16\cos^5\theta}\right)\left(-\frac{c_0\cos^3\theta}{16}\right) - \left(-\frac{c_0^3}{16\cos\theta}\right)^2 = 0. \tag{63}$$

The update

$$\delta s_\rho(z) \approx -\frac{\Gamma_{\kappa\kappa}g_\rho(z) - \Gamma_{\rho\kappa}g_\kappa(z)}{\Gamma_{\kappa\kappa}\Gamma_{\rho\rho} - \Gamma_{\kappa\rho}\Gamma_{\rho\kappa}} \tag{64}$$

is then immediately halted by an undefined step length. This is a powerful internal check.

Correction for cross-talk with the parameter-type approximate Hessian

Finally, let us establish that the parameter-type approximate Hessian and associated quasi-Newton update correctly (to first order) suppresses parameter cross talk. The problem is known to involve both κ and ρ contrasts, so we use two angles of data.

We will work the bulk modulus update example, which requires us to calculate

$$-\frac{\Gamma_{\rho\rho}g_\kappa(z) - \Gamma_{\kappa\rho}g_\rho(z)}{\Gamma_{\kappa\kappa}\Gamma_{\rho\rho} - \Gamma_{\kappa\rho}\Gamma_{\rho\kappa}}. \tag{65}$$

Let us assemble the pieces one at a time. The two numerator quantities are

$$\begin{aligned}\Gamma_{\rho\rho}g_\kappa(z|\theta_1,\theta_2) = &-\frac{c_0^4}{512}\left[\cos^3\theta_1 + \cos^3\theta_2\right]\left[\frac{1}{\cos^5\theta_1} + \frac{1}{\cos^5\theta_2}\right]\left(\frac{\delta s_\kappa}{s_{\kappa_0}}\right)S(z-z_1)\\ &-\frac{c_0^4}{512}\left[\cos^3\theta_1 + \cos^3\theta_2\right]\left[\frac{1}{\cos\theta_1} + \frac{1}{\cos\theta_2}\right]\left(\frac{\delta s_\rho}{s_{\rho_0}}\right)S(z-z_1),\end{aligned} \tag{66}$$

using the gradient in equation (42) and the coefficients in equations (53)-(56) evaluated for the same two angles θ_1 and θ_2, and

$$\begin{aligned}\Gamma_{\kappa\rho}g_\rho(z|\theta_1,\theta_2) = &-\frac{c_0^4}{512}\left[\frac{1}{\cos\theta_1}+\frac{1}{\cos\theta_2}\right]\left[\cos^3\theta_1+\cos^3\theta_2\right]\left(\frac{\delta s_\kappa}{s_{\kappa 0}}\right)S(z-z_1)\\ &-\frac{c_0^4}{512}\left[\frac{1}{\cos\theta_1}+\frac{1}{\cos\theta_2}\right]\left[\frac{1}{\cos\theta_1}+\frac{1}{\cos\theta_2}\right]\left(\frac{\delta s_\rho}{s_{\rho 0}}\right)S(z-z_1),\end{aligned} \tag{67}$$

likewise. In the numerator these two quantities are subtracted. Hence, the most difficult test for robustness against cross-talk, is seen to be passed during the course of the numerator being assembled. Both of these terms have components proportional to δs_ρ, with equal coefficients. So, the differencing in the numerator is sufficient to suppress the cross talk. Now the calculation is proportional to δs_κ only:

$$\begin{aligned}&\Gamma_{\rho\rho}g_\kappa(z|\theta_1,\theta_2) - \Gamma_{\kappa\rho}g_\rho(z|\theta_1,\theta_2)\\ &\quad=\left(\frac{c_0^4}{512}\right)\left[\frac{\cos^3\theta_1}{\cos^5\theta_2}-\frac{2}{\cos\theta_1\cos\theta_2}+\frac{\cos^3\theta_2}{\cos^5\theta_1}\right]\left(\frac{\delta s_\kappa}{s_{\kappa 0}}\right)S(z-z_1).\end{aligned} \tag{68}$$

Meanwhile the determinant in the reciprocal evaluates to

$$\frac{1}{\Gamma_{\kappa\kappa}\Gamma_{\rho\rho}-\Gamma_{\kappa\rho}\Gamma_{\rho\kappa}} = -\left(\frac{c_0^6}{256}\right)^{-1}\left[\frac{\cos^3\theta_1}{\cos^5\theta_2}-\frac{2}{\cos\theta_1\cos\theta_2}+\frac{\cos^3\theta_2}{\cos^5\theta_1}\right]^{-1}, \tag{69}$$

which in a final step suppresses the angle-dependence of the update. Again to first order,

$$-\frac{\Gamma_{\rho\rho}g_\kappa(z)-\Gamma_{\kappa\rho}g_\rho(z)}{\Gamma_{\kappa\kappa}\Gamma_{\rho\rho}-\Gamma_{\kappa\rho}\Gamma_{\rho\kappa}} \approx -\frac{1}{2c_0^2}\left(\frac{\delta s_\kappa}{s_{\kappa 0}}\right)S(z-z_1). \tag{70}$$

The parameter-type quasi-Newton update is equal to the ideal update, therefore, to within a scalar multiple. A second place finish in the presence of multiple parameter variations, the only formula we have considered which accomplishes this.

SUMMARY

We have demonstrated several ways in which a particular form of quasi-Newton update addresses multi-parameter issues in application to pre-critical reflection FWI. These are addressed in a full Newton update, of course, but in the *parameter-type* quasi-Newton update they are addressed in a way which involves almost no additional compuation beyond that done for gradient-based update steps.

The analysis leading to these facts has simultaneously permitted us to parse the behaviour of the elements of Newton and quasi-Newton steps. As a summary aid, we sketch out "maps" of the roles parts of the Newton and quasi-Newton formulas play in effectively updating more than one parameter.

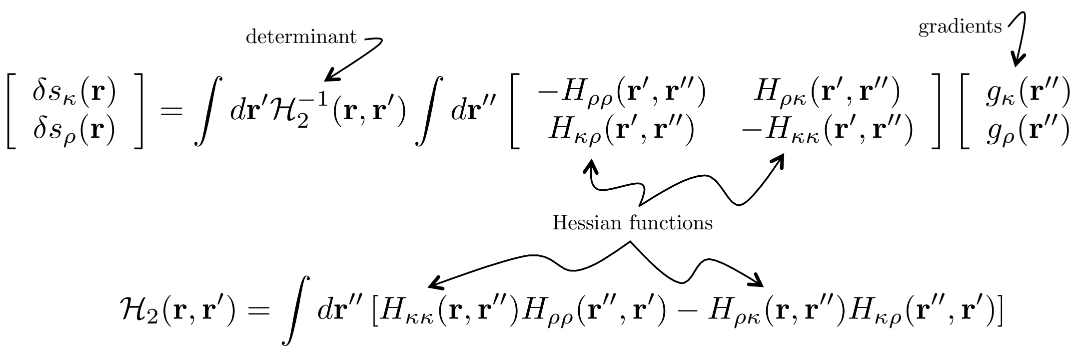
A map of the multiparameter Newton step

In Figure 3, the two-parameter full Newton step formula for acoustic FWI is displayed (the three parameter isotropic-elastic version is provided by Innanen, 2013). It is the hybrid continuous-discrete version of the discrete formula

$$\delta \mathrm{s} = \mathbf{H}^{-1}\mathbf{g}. \tag{71}$$

The two-row column vector at the right end is the gradient vector, and the two integrals containing the combinations of $H_{\kappa\kappa}$, $H_{\kappa\rho}$, $H_{\rho\kappa}$ and $H_{\rho\rho}$ constitute the inverse Hessian. The function $\mathcal{H}_2$ is the generalization of the determinant, and the 2×2 matrix in the middle of the formula is the generalization of the transposed co-factor matrix. As we discussed when we invoked the bivariate template (Innanen, 2013), the determinant/cofactor transpose mixture is analogous to the inverse matrix formula

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \tag{72}$$



$$\begin{bmatrix} \delta s_\kappa(\mathbf{r}) \\ \delta s_\rho(\mathbf{r}) \end{bmatrix} = \int d\mathbf{r}' \mathcal{H}_2^{-1}(\mathbf{r},\mathbf{r}') \int d\mathbf{r}'' \begin{bmatrix} -H_{\rho\rho}(\mathbf{r}',\mathbf{r}'') & H_{\rho\kappa}(\mathbf{r}',\mathbf{r}'') \\ H_{\kappa\rho}(\mathbf{r}',\mathbf{r}'') & -H_{\kappa\kappa}(\mathbf{r}',\mathbf{r}'') \end{bmatrix} \begin{bmatrix} g_\kappa(\mathbf{r}'') \\ g_\rho(\mathbf{r}'') \end{bmatrix}$$

$$\mathcal{H}_2(\mathbf{r},\mathbf{r}') = \int d\mathbf{r}'' \left[H_{\kappa\kappa}(\mathbf{r},\mathbf{r}'')H_{\rho\rho}(\mathbf{r}'',\mathbf{r}') - H_{\rho\kappa}(\mathbf{r},\mathbf{r}'')H_{\kappa\rho}(\mathbf{r}'',\mathbf{r}')\right]$$

FIG. 3. A map of the formula for a two-parameter full Newton step.

A map of the *parameter-type* quasi- Newton step

In fact, by approximating the functions $H_{\kappa\kappa}$, $H_{\kappa\rho}$, $H_{\rho\kappa}$ and $H_{\rho\rho}$ in terms of delta functions, we derive *parameter-type* quasi-Newton step formulas which much more explicitly involve this inverse matrix formula (Figure 4a). For analytic purposes, this requires us to find, in a given example, portions of the full Hessian formulas which correspond to coefficients multiplied by delta functions, which we have done in this paper for the acoustic single interface problem. In practical implementations, this simply requires us to retain the diagonal components of the off-diagonal block matrices of the full Hessian.

Where the important tasks of multiparameter pre-critical reflection FWI are carried out by the parameter-type quasi-Newton step are carried out in the formulas is indicated in Figure 4b.

1. The *determinant* is primarily responsible for correction of the step for angle dependence—balancing, as it were, the update so that to first order the same update is produced for any set of data angles used in the FWI procedure.

2. Inspection of the off-diagonal elements of the transposed cofactor matrix invites a negative interpretation—that, through them, the update in one parameter, say δs_κ, is altered by the ρ gradient, rather than constructed solely using κ. This appears superficially to be a source for cross-talk, in other words. However, it is the opposite. The gradients themselves are already contaminated by cross-talk; these off-diagonal elements in fact subtract off cross-talk noise from the update.

3. The diagonal elements of the cofactor matrix perform a final update balancing, leaving the updates within a scalar multiple of the idealized update.

(a)

$$\begin{bmatrix} \delta s_\kappa(\mathbf{r}) \\ \delta s_\rho(\mathbf{r}) \end{bmatrix} \approx -\overbrace{\frac{1}{\Gamma_{\kappa\kappa}(\mathbf{r})\Gamma_{\rho\rho}(\mathbf{r})-\Gamma_{\rho\kappa}(\mathbf{r})\Gamma_{\kappa\rho}(\mathbf{r})}}^{\text{determinant}} \overbrace{\begin{bmatrix} \Gamma_{\rho\rho}(\mathbf{r}) & -\Gamma_{\rho\kappa}(\mathbf{r}) \\ -\Gamma_{\kappa\rho}(\mathbf{r}) & \Gamma_{\kappa\kappa}(\mathbf{r}) \end{bmatrix}}^{\text{Hessian functions}} \overbrace{\begin{bmatrix} g_\kappa(\mathbf{r}) \\ g_\rho(\mathbf{r}) \end{bmatrix}}^{\text{gradients}}$$

(b)

$$\begin{bmatrix} \delta s_\kappa(\mathbf{r}) \\ \delta s_\rho(\mathbf{r}) \end{bmatrix} \approx -\underbrace{\frac{1}{\Gamma_{\kappa\kappa}(\mathbf{r})\Gamma_{\rho\rho}(\mathbf{r})-\Gamma_{\rho\kappa}(\mathbf{r})\Gamma_{\kappa\rho}(\mathbf{r})}}_{\text{angle-dependence suppression}} \begin{bmatrix} \Gamma_{\rho\rho}(\mathbf{r}) & -\Gamma_{\rho\kappa}(\mathbf{r}) \\ -\Gamma_{\kappa\rho}(\mathbf{r}) & \Gamma_{\kappa\kappa}(\mathbf{r}) \end{bmatrix} \begin{bmatrix} g_\kappa(\mathbf{r}) \\ g_\rho(\mathbf{r}) \end{bmatrix}$$

amplitude correction ($\Gamma_{\rho\rho}(\mathbf{r})$)
cross-talk suppression ($-\Gamma_{\rho\kappa}(\mathbf{r})$)
cross-talk suppression ($-\Gamma_{\kappa\rho}(\mathbf{r})$)
amplitude correction ($\Gamma_{\kappa\kappa}(\mathbf{r})$)

FIG. 4. (a) The components of the parameter-type quasi-Newton update formula. (b) The tasks carried out by the components of the formula, including suppression of angle dependence and update balancing; amplitude correction; cross-talk suppression.

CONCLUSIONS

A criticism made from time to time regarding FWI is that it is a primarily numerical exercise, wherein real wave physics and data analysis questions go unasked as our computers grind away and the answer is (hopefully) converged upon. A bit harsh, perhaps—but containing a grain of truth.

However, FWI is also emerging as a potentially powerful approach for seismic inverse problems, which leverage methods the community already has some awareness of (e.g., RTM, incorporation of well control, wavelet estimation), and in which prior information, regularization, and normed data misfits, can all be incorporated relatively seamlessly.

Part of the heavy lifting to be done involves finding ways not only of making FWI algorithms, but making FWI algorithms which are intelligible, and whose components are easily exposed to analysis. And, making the analysis methods by which they are characterized.

This paper has concerned one aspect of this large problem, namely understanding what the roles of the quantities familiar in FWI (e.g., gradient and Hessian) are in incorporating the pre-critical reflection amplitude information we normally subject to AVO analysis. We know almost instinctively, from years of community progress in AVO inversion, how angle variations in the reflection coefficient can be combined in sums and differences to separate out the influence of various parameters—to suppress cross-talk, the FWI practitioner would say. The purpose of this paper has been to figure out how to point to parts of the FWI Newton and quasi-Newton update mathematics, and say, "Here is where cross-talk is managed".

ACKNOWLEDGMENTS

This work was funded by the CREWES project. CREWES sponsors and research personnel are thanked.

REFERENCES

Castagna, J. P., and Backus, M., 1993, Offset-dependent reflectivity: theory and practice of AVO analysis: SEG.

Castagna, J. P., Swan, H. W., and Foster, D. J., 1998, Framework for AVO gradient and intercept interpretation: Geophysics, **63**, No. 3, 948–956.

Clayton, R. W., and Stolt, R. H., 1981, A Born-WKBJ inversion method for acoustic reflection data: Geophysics, **46**, No. 11, 1559–1567.

Foster, D. J., Keys, R. G., and Lane, F. D., 2010, Intepretation of AVO anomalies: Geophysics, **75**, 75A3–75A13.

Innanen, K. A., 2013, A framework for iterative multiparameter seismic inversion: CREWES Annual Research Report, **25**.

Innanen, K. A., and Weglein, A. B., 2007, On the construction of an absorptive-dispersive medium model via direct linear inversion of reflected seismic primaries: Inverse Problems, **23**, 2289–2310.

Raz, S., 1981, Direct reconstruction of velocity and density profiles from scattered field data: Geophysics, **46**, 832.

Stolt, R. H., and Weglein, A. B., 1985, Migration and inversion of seismic data: Geophysics, **50**, No. 12, 2458–2472.

Weglein, A. B., Araújo, F. V., Carvalho, P. M., Stolt, R. H., Matson, K. H., Coates, R. T., Corrigan, D., Foster, D. J., Shaw, S. A., and Zhang, H., 2003, Inverse scattering series and seismic exploration: Inverse Problems, R27–R83.

Weglein, A. B., Violette, P. B., and Keho, T. H., 1986, Using multiparameter Born theory to obtain certain exact multiparameter inversion goals: Geophysics, **51**, 1069–1074.