Introduction

Seismic full waveform inversion (FWI), while making significant progress in many interrelated domains (e.g., Plessix et al., 2013; Operto et al., 2013; Brossier et al., 2013), has not as yet given rise to a fully realized methodology for precritical reflection data (Kelly et al., 2013). However, on the strength of evidence that acquisition and well-control may provide some keys to making precritical reflection FWI practical (e.g., Margrave et al., 2012a,b), we are considering a set of outstanding theoretical questions about how it might behave. Central to this—if we anticipate using the information contained in precritical amplitudes—is the development of a clear theoretical reconciliation between FWI and amplitu-variation-with-offset (AVO) inversion. With a bridge between FWI and AVO in place, the reach of FWI, supported by decades of community experience in AVO, would be significantly extended.

In FWI it is understood that the inverse Hessian operator must correct many of the amplitude-related errors, for instance cross-talk, to which gradient-based methods are exposed (Operto et al., 2013). AVO inversion methods (Hampson and Russell, 1990) and their inverse scattering generalizations (e.g., Clayton and Stolt, 1981) are, through weighted differences of reflection coefficients as functions of angle, protected to first order from cross-talk. Does a properly assembled full Newton update in FWI do the same thing automatically? Where in the inverse Hessian operator are the parts responsible for AVO-like cross-talk suppression? In two communications on the subject, of which this is the first, we report on our efforts to answer these questions. A mixed discrete-continuous form for a two parameter acoustic full inversion methods (Hampson and Russell, 1990) and their inverse scattering generalizations (e.g., Clay-

Continuous-discrete formulation of FWI – acoustic (bulk modulus $\kappa$ and density $\rho$) case

We seek the functions $s_{\kappa}(r) = \kappa^{-1}(r)$ and $s_{\rho}(r) = \rho^{-1}(r)$ which minimize an objective functional $\Phi(s_{\kappa}, s_{\rho})$. A Newton step $(\delta s_{\kappa}, \delta s_{\rho})^T$ towards this minimum is related to the gradient $(g_{\kappa}, g_{\rho})^T = (\partial \Phi / \partial s_{\kappa}, \partial \Phi / \partial s_{\rho})^T$ by

$$
\begin{bmatrix}
g_{\kappa}(r) \\
g_{\rho}(r)
\end{bmatrix} = - \int dr' \begin{bmatrix} H_{kk}(r, r') & H_{kp}(r, r') \\ H_{pk}(r, r') & H_{pp}(r, r') \end{bmatrix} \begin{bmatrix} \delta s_{\kappa}(r') \\ \delta s_{\rho}(r') \end{bmatrix},
$$

(1)

where, for this two parameter problem, the Hessian consists of four block operators with the kernels

$$
H_{AB}(r, r') = \frac{\partial^2 \Phi(s_{\kappa}, s_{\rho})}{\partial s_A(r) \partial s_B(r')},
$$

(2)

For special cases in which the block operators commute, the inverse of equation (1) can be written

$$
\begin{bmatrix}
\delta s_{\kappa}(r) \\ \delta s_{\rho}(r)
\end{bmatrix} = - \int dr' H^{-1}(r, r') \int dr'' \begin{bmatrix} H_{pp}(r'', r') & -H_{kp}(r'', r') \\ -H_{pk}(r'', r') & H_{kk}(r'', r') \end{bmatrix} \begin{bmatrix} g_{\kappa}(r'') \\ g_{\rho}(r'') \end{bmatrix},
$$

(3)

where $H^{-1}$ is the inverse of

$$
H(r, r') = \int dr'' \left[ H_{kk}(r, r'') H_{pp}(r'', r') - H_{kp}(r, r'') H_{pk}(r'', r') \right]
$$

(4)

in the sense that $\int dr'' H^{-1}(r, r'') H(r'', r') = \delta(r - r')$. In a gradient-based update the inverse Hessian is replaced by a diagonal operator with two scalar unknowns, which are to be determined through a
line search. For rhetorical purposes let us reduce equation (3) to one of gradient type in two stages. Let us first suppose that the inverse Hessian is not itself diagonal, but that its four block operators are. That is, we approximate \( H_{AB}(\mathbf{r}, \mathbf{r}') \approx h_{AB}(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}') \). Thus the update at one point in space will no longer be sensitive to gradient values at other points in space. The generalized determinant becomes \( \mathcal{H}(\mathbf{r}, \mathbf{r}') \approx [h_{kk}(\mathbf{r})h_{pp}(\mathbf{r}) - h_{kp}(\mathbf{r})h_{pk}(\mathbf{r})] \delta(\mathbf{r} - \mathbf{r}') \), and the update reduces to

\[
\begin{bmatrix}
δs_κ(\mathbf{r}) \\
δs_ρ(\mathbf{r})
\end{bmatrix}
≈
\begin{bmatrix}
\frac{1}{h_{kk}(\mathbf{r})h_{pp}(\mathbf{r}) - h_{kp}(\mathbf{r})h_{pk}(\mathbf{r})}
\end{bmatrix}
\begin{bmatrix}
h_{pp}(\mathbf{r}) & -h_{kp}(\mathbf{r}) \\
-h_{pk}(\mathbf{r}) & h_{kk}(\mathbf{r})
\end{bmatrix}
\begin{bmatrix}
g_κ(\mathbf{r}) \\
g_ρ(\mathbf{r})
\end{bmatrix}.
\]

(5)

We will refer to this as a parameter-type update. A bulk modulus update of parameter type can be influenced by the density gradient, and vice versa, because of remaining off-diagonal terms in the inverse Hessian approximation. To reduce from a parameter-type to a gradient-based update, we suppress these:

\[
h_{kk}(\mathbf{r}) \approx -μ_κ^{-1}, \quad h_{kp}(\mathbf{r}) \approx 0, \quad h_{pk}(\mathbf{r}) \approx 0, \quad h_{pp}(\mathbf{r}) \approx -μ_ρ^{-1},
\]

which leads to the gradient-based update

\[
\begin{bmatrix}
δs_κ(\mathbf{r}) \\
δs_ρ(\mathbf{r})
\end{bmatrix}
≈
\begin{bmatrix}
μ_κg_κ(\mathbf{r}) \\
μ_ρg_ρ(\mathbf{r})
\end{bmatrix}.
\]

(7)

Analytic reconstruction of a single acoustic boundary

If we were to make algorithms out of the formulas in equations (3), then (5), and then (7), we would observe in that progression a decrease in computational cost, but also a decrease in the sophistication with which data are interrogated—for instance, given precritical reflection data, in their “intelligent” use of AVO information. We will now quantitate this with an analytic example. Let a plane wave impinge on a boundary separating a medium with properties \( s_0 = 1/κ_0 \) and \( s_0 = 1/ρ_0 \) from one with \( s_1 = 1/κ \) and \( s_1 = 1/ρ \), where \( κ \) and \( ρ \) are the bulk modulus and density respectively. Defining \( δs_κ = s_κ - s_κ_0 \) and \( δs_ρ = s_ρ - s_ρ_0 \), the associated reflection coefficient is, to first order in \( δs_κ \), \( δs_ρ \), and \( sin^2θ \),

\[
R(θ) ≈ \frac{1}{4} \frac{1}{cos^2θ} \left( \frac{δs_κ}{s_κ_0} \right) - \frac{1}{4} \frac{1}{cos^2θ} \left( \frac{δs_ρ}{s_ρ_0} \right).
\]

(8)

In a FWI problem, if the initial or background medium were homogeneous, and the data were due to a reflection from a single acoustic interface at depth \( z_1 \), then the ideal updates, i.e., those which take us in a single step to the correct model, would be

\[
δs_κ(z) = δs_κS(z - z_1), \quad \text{and} \quad δs_ρ(z) = δs_ρS(z - z_1),
\]

(9)

where \( S \) is the Heaviside function with step at \( z_1 \), and where \( δs_κ \) and \( δs_ρ \) satisfy equation (8).

![Figure 1](https://example.com/figure1.png)

**Figure 1** (a) \( θ, ω, k_κ \) and \( q_κ \) relations. (b) Reflection from an acoustic interface.

We construct and assemble the ingredients for a FWI gradient under these same circumstances, and compare the resulting update to this benchmark. We assume (i) that the data are measurements of an
acoustic wave field $P(k_g, z, x, z_s, \omega)$ and (ii) that during the nth FWI iteration we have access to the modelled field $G_n(k_g, z, x, z_s, \omega)$. For instance at $n = 0$ we have (e.g., Clayton and Stolt, 1981)

$$G_0(k_g, z, x, z_s, \omega) = \rho_0 e^{-ik_g z} \frac{e^{i\omega(|z-z_s|)}}{2\pi q_g}, \quad G_0(x_g, z, x, z_s, \omega) = \frac{\rho_0}{2\pi} \int dk' e^{ik'(x-x_g)} \frac{e^{i\omega(|z-z_s|)}}{2k'}.$$  \hspace{1cm} (10)

where $q_g = (\omega/c_0)(1 - k_g^2 c_0^2/\omega^2)^{1/2}$ and $q' = (\omega/c_0)(1 - k'_g c_0^2/\omega^2)^{1/2}$, and where $c_0 = \sqrt{k_0/\rho_0} = \sqrt{s_{bg}/s_{bg}}$ is the constant background acoustic wave velocity (see Figure 1a). We next must quantify the sensitivity of the $G_n$ to small changes in the depth-variability of $\kappa$ and $\rho$. These can be shown to be

$$\frac{\partial G_n}{\partial \kappa_n(z)} = -\omega^2 \int dx' G_n(k_g, z_g, x', z) G_n(x', z, x, z_s)$$  \hspace{1cm} (11)

and

$$\frac{\partial G_n}{\partial \rho_n(z)} = \int dx' \left[ \frac{\partial G_n(k_g, z_g, x', z)}{\partial z} \frac{\partial G_n(x', z_s, z_s)}{\partial z} - G_n(k_g, z_g, x', z) \frac{\partial^2 G_n(x', z_s, z_s)}{\partial x^2} \right].$$  \hspace{1cm} (12)

Finally, if the data are due to a reflection from a single interface (Figure 1b), the complex conjugates of the residuals are

$$\delta P_n^\ast(k_g, \omega) = -\rho_0 R(\theta) e^{-i2q_g z_1}.$$  \hspace{1cm} (13)

If we select $\Phi(s_k, s_p) = \frac{1}{2} \sum_k \int d\omega |\delta P|^2$, the gradients $g_{\kappa}$ and $g_{\rho}$ at the nth iteration can be shown to involve combinations of the sensitivities in equations (11)–(12) and the nth residuals $\delta P$: \hspace{1cm}

$$g_{\kappa}^{(n)}(z) = \sum_k \int d\omega \omega^2 \int dx' G_n(k_g, z_g, x', z, \omega) G_n(x', z_s, z_s, \omega) \delta P_n^\ast,$$

$$g_{\rho}^{(n)}(z) = \sum_k \int d\omega \int dx' \left[ G_n(k_g, z_g, x', z) \frac{\partial^2 G_n(x', z_s, z_s)}{\partial x^2} - G_n(k_g, z_g, x', z) \frac{\partial G_n(x', z_s, z_s)}{\partial z} \right] \delta P_n^\ast.$$  \hspace{1cm} (14)

To study the first FWI iteration ($n = 0$) with an eye for AVO-like behaviour, we substitute equations (10) into these gradient formulas, and then transform from $k_g$ to incidence angle $\theta$ (as per Figure 1a). For small angles ($\tan^2 \theta \approx \sin^2 \theta$) we thus obtain

$$g_{\kappa}^{(0)}(z) = -\frac{\rho_0^2 c_0^2}{4} \sum_\theta \int d\omega \ e^{2iq_g z} \left( \frac{1}{\cos^2 \theta} \right) \delta P_0^\ast,$$

$$g_{\rho}^{(0)}(z) = -\frac{\rho_0^2}{4} \sum_\theta \int d\omega \ e^{2iq_g z} \left( \cos^2 \theta \right) \delta P_0^\ast.$$  \hspace{1cm} (15)

To then consider this update given a reflection from a single interface, we substitute equation (13) into (14), change integration variables from $\omega$ to $2q_g$, and integrate. Given $N$ angles, we have

$$g_{\kappa}^{(0)}(z|\theta_1, \ldots, \theta_N) = \frac{c_0 \rho_0^3 \pi}{4} \left[ \sum_{j=1}^N \frac{R(\theta_j)}{\cos^3 \theta_j} \right] S(z-z_1),$$

$$g_{\rho}^{(0)}(z|\theta_1, \ldots, \theta_N) = \frac{c_0 \rho_0^3 \pi}{4} \left[ \sum_{j=1}^N \frac{R(\theta_j) \cos \theta_j}{\cos \theta_j} \right] S(z-z_1).$$  \hspace{1cm} (15)

AVO inverse problems in $m$ parameters are generally well-posed given at least $m$ angles of data. The $N = 2$ case of equation (15), with the inclusion of two scalar unknowns to be fixed by a line search, therefore forms a well-posed gradient-based update $(\delta g_{\kappa}^{(0)}, \delta g_{\rho}^{(0)})^T$. To analyze it, we substitute for each
instance of $R(\theta)$ equation (8), producing, finally, expressions for the proposed gradient-based updates in terms of the idealized updates:

$$
\delta s_k^b(z) = \mu_{k_0} g_k^{(0)}(z|\theta_1, \theta_2) \approx -\mu_{k_0} \left(\frac{c_0^3 \rho_0^3 \pi}{16} \right) \left[ \frac{1}{\cos^3 \theta_1} \right] \left( \frac{\delta s_k}{s_{k_0}} \right) S(z-z_1)
$$

$$
- \mu_{k_0} \left(\frac{c_0^3 \rho_0^3 \pi}{16} \right) \left[ \frac{1}{\cos \theta_1} + \frac{1}{\cos \theta_2} \right] \left( \delta s_p \right) S(z-z_1),
$$

$$
\delta s_p^b(z) = \mu_{p_0} g_p^{(0)}(z|\theta_1, \theta_2) \approx -\mu_{p_0} \left(\frac{c_0^3 \rho_0^3 \pi}{16} \right) \left[ \cos^3 \theta_1 + \cos^3 \theta_2 \right] \left( \delta s_p \right) S(z-z_1)
$$

$$
- \mu_{p_0} \left(\frac{c_0^3 \rho_0^3 \pi}{16} \right) \left[ \frac{1}{\cos \theta_1} + \frac{1}{\cos \theta_2} \right] \left( \delta s_k \right) S(z-z_1).
$$

(16)

Criticism of gradient-based methods for precritical reflection data

Equations (15)-(16) can now be used to expose some issues in precritical FWI. These can either be seen as criticisms of gradient-based updates, or as indications of what we must seek in approximating the inverse Hessian. 1. **Response to ill-posedness.** Nothing stops equation (15) from being used to form updates in the case $N = 1$, i.e., using a single angle of data, though from AVO theory and practice this problem is known to be ill-posed. Insufficiency of data is not internally diagnosed by gradient-based methods. 2. **Angle-dependent gradients.** The gradients for bulk modulus and density depend explicitly on the input data angles. This puts an extra burden on the line-search, which must determine $\mu_{k_0}$ and $\mu_{p_0}$ to accommodate this variation. 3. **Cross-talk.** Most critically, from equation (16) the gradient-based updates are seen to depend explicitly on both gradients, which means cross-talk will be a key problem in precritical FWI. No scalar $\mu_{k_0}$ and $\mu_{p_0}$ pair can be determined which suppresses this problem.

Conclusions

Precritical reflection FWI gradients, expressed in the above forms, though on their own evidently deficient as updates, can now be used to analyze more sophisticated updates—e.g., the parameter-type update in equation (5). It is possible to show that these elements of the inverse Hessian correct the three problems detailed above, and produce a FWI update which generalizes direct AVO inversion. The question of how to include these tractably in an already computationally intensive FWI update remains open. Here our purpose has been to develop analytical methods to tell us precisely which parts of the inverse Hessian it is worth attempting to include, as we vie with the computational aspect.

Acknowledgments

This research was funded by the CREWES Project and NSERC.

References


