# Residual dependent FWI sensitivities based on direct nonlinear inverse scattering

Kris Innanen

## ABSTRACT

Nonlinear, residual dependent FWI sensitivities are the outgrowth of the observation that certain direct nonlinear procedures, available in special case environments/schemes, such as AVO inversion or 1D/1.5D direct nonlinear inverse scattering imaging and inversion, have no FWI generalization – they cannot be found as special cases of any standard FWI procedure. In order to derive general FWI schemes that reduce in this way, an extension of some kind is required to our ideas of FWI sensitivity. A proposal from 2014, namely that the sensitivities be created by varying not the current model iteration, but the forthcoming iteration, is pursued in this paper to confirm that this approach correctly reduces to an existing 1D normal incidence direct nonlinear imaging and inversion scheme, previously derived from inverse scattering considerations.

## INTRODUCTION

The purpose of this paper is to refine and expand on some ideas discussed in the 2014 CREWES report (Innanen, 2014b), on the extension of full waveform inversion sensitivities to incorporate nonlinearity. Specifically, extensions such that techniques available from direct nonlinear inverse scattering (e.g., Weglein et al., 1981; Stolt and Jacobs, 1980; Weglein et al., 2003; Shaw et al., 2004; Innanen, 2008; Zhang and Weglein, 2009a,b) can be merged with the ideas and concepts of full waveform inversion.

The basic framework will be reviewed and refined in this paper, and we will add a second case to the reflectivity examples incorporated previously. In the remainder of this introduction, the problem that is solved with nonlinear sensitivities will be laid out in some detail. Following this, the basic equations of scalar multidimensional FWI are discussed, and the general nonlinear sensitivities idea is expressed first qualitatively, and second quantitatively. Then we treat two special cases, showing that FWI updates reduce in these special cases to forms which have been shown elsewhere to incorporate nonlinearity in a way which adds significant differential benefit.

This research is "blue sky" — it will become of importance when FWI overcomes current hurdles in making multiparameter inversion computable and stable, in making algorithms for which bandlimited data involving all surface and body waves, with all reflected, refracted, and diving modes, are acceptable input. However, when that happens, and our interests begin to extend to determination of, e.g., anisotropic parameters from backscattered data and many incidence and azimuth angles, incorporating nonlinearity within updates may become a necessity. Preparing techniques in advance of this need is an important task.

## "Missing" full waveform inversion procedures

The motivation for the work in this paper seems to be best expressed by pointing to "missing" FWI procedures — ones which we wish existed, but don't. The idea of a missing FWI procedure is a corollary of the existence of FWI forms that agree with recognizable, pre-existing inverse procedures in some simplified limit. An example, discussed by Innanen (2014a), is the particular formulation of multiparameter FWI, which, in the special case of reconstruction of a single interface from reflection data, was found to reduce to an iterative type of standard linearized amplitude-variation-with-offset (AVO) inversion.

Having made this analysis, it is possible to say: "There exists a particular FWI procedure which generalizes iterative linear AVO inversion." Now it will be convenient to re-state this fact in a backward-sounding way: "Iterative linear AVO inversion *has no missing FWI generalization*." In this paper we will be interested in cases in which there *are* missing FWI procedures. That is, powerful formulations of inversion which exist in special cases, but which have no FWI generalization.

## Example: iterative nonlinear AVO inversion

An example can be found in the same single-interface/reflection environment in which linearized AVO inversion is derived. Suppose surface reflection data were acquired somewhere above a single horizontal interface separating a scalar P-wave velocity of $c_0 = 1500$m/s in the upper half-space from a velocity of $c_1 = 2200$m/s in the lower half space. From these data, the reflection coefficient $R(\theta)$ is extracted for a range of angles. The problem is, given $c_0$ and measurements of the coefficient $R(\theta)$, to determine $c_1$.

The problem can be approached in the following way. Parameterizing in terms of squared slowness, we let $s = c^{-2}$. The jump from $c_0$ to $c_1$ across the boundary is expressible in these terms as

$$\left(\frac{\Delta s}{s_0}\right) = \frac{s_1 - s_0}{s_0},\tag{1}$$

and the reflection coefficient can be expressed as a series in orders of this perturbation and $\sin^2\theta$ as

$$\begin{aligned}
R(\theta) &= -\frac{1}{4}\left(1 + \sin^2\theta + ...\right)\left(\frac{\Delta s}{s_0}\right) + \left(\frac{1}{8} + \frac{1}{4}\sin^2\theta + ...\right)\left(\frac{\Delta s}{s_0}\right)^2 - ... \\
&\approx -\frac{1}{4}\left(1 + \sin^2\theta\right)\left(\frac{\Delta s}{s_0}\right).
\end{aligned}\tag{2}$$

In the second line we have linearized $R$ in both $(\Delta s/s_0)$ and $\sin^2\theta$. The linearization may suffer from significant truncation error, in fact the roughly $50\%$ jump in P-wave velocities across the boundary was chosen to cause just such a problem. By comparing plots of the exact reflection coefficient vs. the linearization (Figure 1) we confirm that the *nonlinearity* of the medium property/wave field relationship, contained in the terms left out of the second line of equation (2), will be significant here.
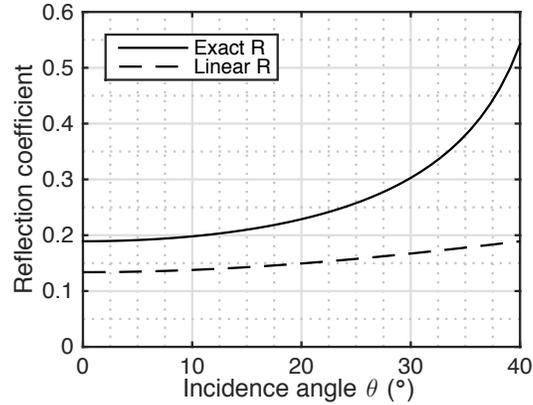
FIG. 1. Comparison between the exact and linearized reflection coefficients for a 50% contrast. The discrepancy is a strong indication that nonlinearity will be a factor in analyzing this reflectivity.
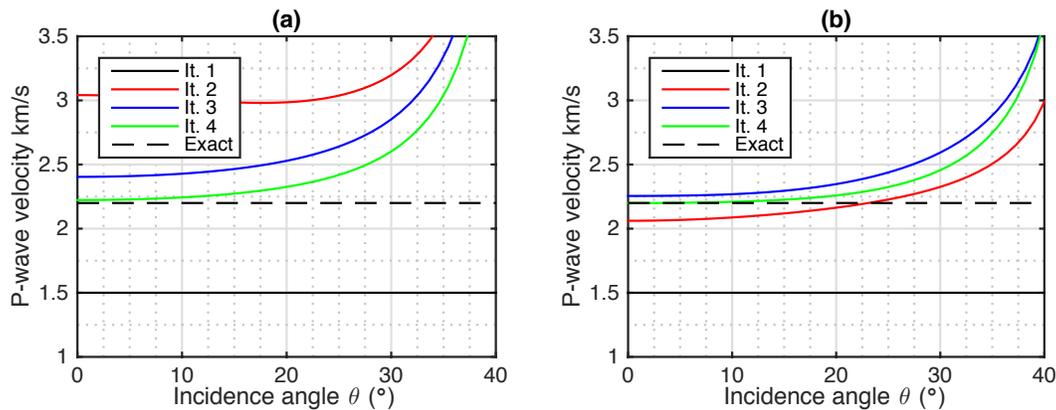


FIG. 2. Iterative reconstruction of the boundary P-wave velocity variation. (a) Iterative linear; (b) iterative nonlinear (second order). Iterations 1–4 included.

Solving the second line of equation (2) for $(\Delta s / s_0)$ in terms of $R$, i.e., direct linear inversion, will likely not work too well, in light of what we see in Figure 1. However, nonlinearity can be accommodated in an otherwise linear scheme through iteration. An iterative linear AVO inversion procedure could be devised to determine $c_1$ as follows. At each angle $\theta$ where a datum $R$ is available, the second line of equation (2) is used to determine $(\Delta s / s_0)$. The $\Delta s$ in the numerator is added to $s_0$ to update $s_1 = s_0 + \Delta s$. This updated guess is then used to generate a modelled $R$, which is subtracted from the observed reflection coefficient to create a residual $\Delta R$. The second line in equation (2) is again used to determine $\Delta s$, this time the change arising from a left hand side of $\Delta R$. This updating is then repeated. The result is displayed in Figure 2a: the first guess is the constant 1.5km/s, plotted as a solid black line, then, iterating, we jump to the red line, then the blue line, then the green line. So, in spite of the large linearization error, we observe a tendency to converge towards the right answer of 2.2km/s (which is plotted as a dashed line), especially at the lower angles.
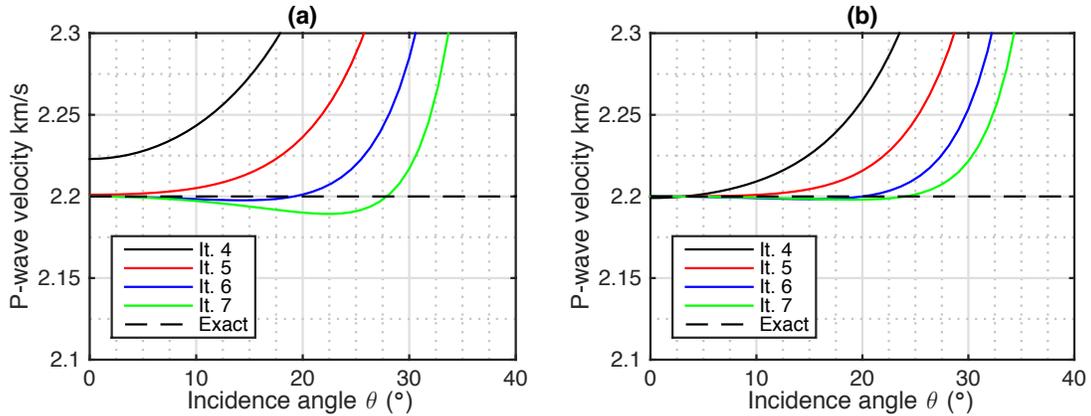
FIG. 3. Iterative reconstruction of the boundary P-wave velocity variation. (a) Iterative linear; (b) iterative nonlinear (second order). Iterations 4–7 included.

A different way to accommodate nonlinearity is to do so directly, that is, by truncating the series in the first line of equation (2) at some order higher than first, and then solving for $(\Delta s/s_0)$ in terms of $R$. This can be done for instance via the technique of series reversion (Abramowitz and Stegun, 1972), as used by Innanen (2011). The results coincide at low order with those derived from direct nonlinear inverse scattering considerations in the acoustic, single parameter limit (Zhang, 2006). Generally there will be remaining truncation error, but, this can be mitigated by iteration, same as in the linear case. The result of doing so for an inverse series truncated after second order is illustrated in Figure 2b. Again, the first iteration is the initial guess (the black line at 1.5km/s), and the iterations proceed to the red, blue, and then green lines. Comparing the iterative second order procedure against the iterative linear procedure (Figure 2b vs 2a), the former suggests a significant improvement in convergence rate, and convergence over a larger angle range.

If we iterate further (Figure 3), it also appears that the low order nonlinear result is more stable – by iteration 7 the iterative linear result (Figure 3a) has taken on a pathology in which the lower angles are incorrectly low as convergence is sought at higher angles. No such deviation is seen in the 2nd order iterative nonlinear scheme Figure 3b. So in at least some idealized circumstances, an iterative nonlinear AVO inversion scheme shows very significant differential benefit over iterative linear.

In order to take advantage of these features, but not be restricted to an AVO environment, a natural impulse is to seek whatever FWI scheme generalizes iterative nonlinear AVO inversion. Such a procedure would presumably treat backscattered seismic amplitudes with a similar recognition of the valuable role nonlinearity might play, but on the more general backdrop of the full volume scattering wave physics model. Because direct nonlinear inverse scattering (Weglein et al., 2003; Shaw et al., 2004; Zhang and Weglein, 2009a) generalizes this sort of direct nonlinear AVO inversion, but in a non-iterative manner, the sought after FWI scheme could alternatively be thought of as providing an iterative generalization of direct nonlinear inverse scattering formulations.

Unfortunately, the generalization of nonlinear iterated AVO inversion is an example of a

*missing FWI procedure.* No variant of Newton update, which simultaneously maintains the basic definition of its ingredients (e.g., sensitivities), and seeks to minimize the sum of the squared differences between measured and predicted data, reduces in the "single horizontal interface limit" to iterative nonlinear AVO inversion.

## Options

If standard descent based theories cannot provide the generalization we seek, and we are bent on finding such a generalization, we have several options:

1. Seek to minimize a more complex objective function;

2. Investigate global, as opposed to local, inversion methods; or,

3. Navigate the objective function with updates from an altered sensitivity calculation.

It is not clear which of these options is optimal. The approach we will pursue in this paper, the third, in some respects seems like the worst of the three, because if we alter the basic definitions of any well characterized theoretical framework, it is difficult to see what all of the consequences will be, even if the immediate consequences are positive. We have gone this route because of the familiar, though altered, quantities which arise, and our ability to interpret them in the framework of seismic processing. Evidence of the validity of the approach takes the form of analysis of two independent reduced cases. The first is a recapitulation of the nonlinear AVO analysis presented in 2014; the second is an extension of the idea to the case of direct nonlinear impedance inversion. In both cases, the same well defined procedure for creating nonlinear sensitivities produces existing procedures.

Still, we must emphasize that the situation this theory is in right now is: a certain alteration of the definition of the sensitivities solves some of the immediate problems detailed above, but it is far from certain that other difficulties, so far invisible, will not appear.

## WAVE EQUATIONS AND BASIC FWI QUANTITIES

Let us first set out the basic equations by which the standard and nonstandard FWI procedures will be constructed.

## Equations

The field $P$ due to a point source at $\mathbf{r}_s$, and measured at point $\mathbf{r}$, in the actual medium, will be assumed to satisfy

$$\left[\nabla^2 + \omega^2 s(\mathbf{r})\right] P(\mathbf{r}, \mathbf{r}_s) = \delta(\mathbf{r} - \mathbf{r}_s), \tag{3}$$

where $\omega$ is the angular frequency, and where $s = c^{-2}$ is the squared reciprocal of the actual P-wave velocity. At any given iteration of FWI, the modeled field $G_n(\mathbf{r}, \mathbf{r}_s) = G(\mathbf{r}, \mathbf{r}_s | s_n)$ will be assumed to be available analytically or through simulation, and to satisfy the same

equation but in the presence of provisional model $s_n$:

$$\left[\nabla^2 + \omega^2 s_n(\mathbf{r})\right] G(\mathbf{r}, \mathbf{r}_s | s_n) = \delta(\mathbf{r} - \mathbf{r}_s). \tag{4}$$

The field $P$ will be available on a measurement surface remote from the regions of interest.

**Objective function, gradient and approximate Hessian**

Data are measurements of $P$ at points $\mathbf{r} = \mathbf{r}_g$ on a measurement surface, and these $P(\mathbf{r}_g, \mathbf{r}_s)$ can be compared to the $G(\mathbf{r}_g, \mathbf{r}_s | s_n)$. The objective function, which involves the residuals $\delta P(\mathbf{r}_g, \mathbf{r}_s | s_n) = P(\mathbf{r}_g, \mathbf{r}_s) - G(\mathbf{r}_g, \mathbf{r}_s | s_n)$:

$$\phi(s_n) = \frac{1}{2} \sum_{s,g} \int d\omega |\delta P(\mathbf{r}_g, \mathbf{r}_s | s_n)|^2, \tag{5}$$

is minimized iteratively, during which a model iterate $s_n(\mathbf{r})$ is modified by the update $\delta s_n(\mathbf{r})$ in order to determine $s_{n+1}(\mathbf{r}) = s_n(\mathbf{r}) + \delta s_n(\mathbf{r})$; unindexed $\delta s(\mathbf{r})$ represents a general variation in the medium. A Gauss-Newton update involving $\phi(s_n)$ has the form

$$\delta s_n(\mathbf{r}) = -\int d\mathbf{r}' H_n^{-1}(\mathbf{r}, \mathbf{r}') g_n(\mathbf{r}'), \tag{6}$$

where $g_n$ is the gradient and $H_n^{-1}$ is an approximation to the inverse Hessian, respectively

$$
\begin{aligned}
g_n(\mathbf{r}) &= -\sum_{g,s} \int d\omega \frac{\partial G(\mathbf{r}_g, \mathbf{r}_s | s_n)}{\partial s_n(\mathbf{r})} \delta P^*(\mathbf{r}_g, \mathbf{r}_s | s_n), \quad \text{and} \\
H_n(\mathbf{r}, \mathbf{r}') &= \sum_{g,s} \int d\omega \frac{\partial G(\mathbf{r}_g, \mathbf{r}_s | s_n)}{\partial s_n(\mathbf{r})} \frac{\partial G^*(\mathbf{r}_g, \mathbf{r}_s | s_n)}{\partial s_n(\mathbf{r}')},
\end{aligned}
\tag{7}
$$

the former involving the complex conjugate of the residuals. The Fréchet derivatives $\partial G / \partial s_n$, also known as the sensitivities (the term we will use in this paper), provide the weights needed to alter the direction in model space defined by the residuals such that it points in the direction of steepest descent of $\phi$. Our aim in this paper is to alter the sensitivities, guided by nonlinear direct inverse scattering, in such a way that the resulting FWI update accommodates one or more of a range of types of nonlinearity in the relationship between medium properties and backscattered wave amplitudes.

## GENERAL PROCEDURE FOR COMPUTING NONLINEAR SENSITIVITIES

The ideas behind the general procedure for computing sensitivities are unusual, but they have an intuitive appeal, especially for those with some experience with scattering diagrams (Clayton and Stolt, 1981; Weglein et al., 1997, 2003; Malcolm and de Hoop, 2005; Innanen, 2009). To develop the intuitive side of the derivations a bit, we start with a partially qualitative discussion.
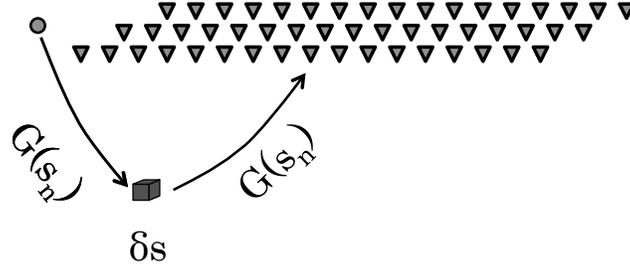
FIG. 4. Illustration of the perturbation in the wave field caused by a local perturbation in the $s_n$th medium.

## Qualitative development

Let $G(s_n)$ represent the wave propagating in medium $s_n$. If the medium is varied locally by the amount $\delta s$, such that the field varies by the amount $\delta G(s_n)$, the sensitivity appears as the coefficient $A$ in the expansion

$$\delta G(s_n) = A\delta s + ... \quad . \tag{8}$$

In Figure 4 the character of the perturbed field in the presence of the perturbed medium is illustrated. The circle represents the seismic source, which causes a wave response in the receivers, represented by triangles. A solid arrow represents the wave propagating in $s_n$. This is a convenient representation, but it gives the impression of a wave propagating in a smoothly varying medium, which is not always the case. We emphasize that the arrows represent all wave processes occurring between its ends, which may be quite a complex set as iterations progress in FWI. In any case, the perturbation $\delta s$ gives rise to a field variation involving the single scattering interaction depicted in the Figure.

## Variation in terms of the $n + 1$th medium iterate

To formulate nonlinear sensitivities calculation, we make a single alteration to the definition of the sensitivities, and then let that change play out. We vary not $G(s_n)$ but rather the field $G(s_{n+1})$, where $s_{n+1}(\mathbf{r})$ is the model iterate we are in the process of determining, and setting as our revised sensitivities $A'$ where

$$\delta G(s_{n+1}) = A'\delta s + ... \quad . \tag{9}$$

This will produce a sensitivity which mediates propagation to and from the variation point $\delta s$, but also introduce propagations between the source, receiver, variation point, and all regions of the medium involving model residuals $\Delta s_n(\mathbf{r}) = s_{n+1}(\mathbf{r}) - s_n(\mathbf{r})$. The qualitative picture is illustrated in Figure 5a. To construct the expansion in equation (13), we construct two series, beginning with equations (3)–(4). The first series is the field in the $n + 1$th medium expanded about the field in the $n$th medium:

$$G(\mathbf{r}_g, \mathbf{r}_s|s_{n+1}) = G(\mathbf{r}_g, \mathbf{r}_s|s_n) - \omega^2 \int d\mathbf{r}' G(\mathbf{r}_g, \mathbf{r}'|s_n)\Delta s_n(\mathbf{r}')G(\mathbf{r}', \mathbf{r}_s|s_n)$$
$$+ \omega^4 \int d\mathbf{r}' G(\mathbf{r}_g, \mathbf{r}'|s_n)\Delta s_n(\mathbf{r}') \int d\mathbf{r}'' G(\mathbf{r}', \mathbf{r}''|s_n)\Delta s_n(\mathbf{r}'')G(\mathbf{r}'', \mathbf{r}_s|s_n) + ..., \tag{10}$$
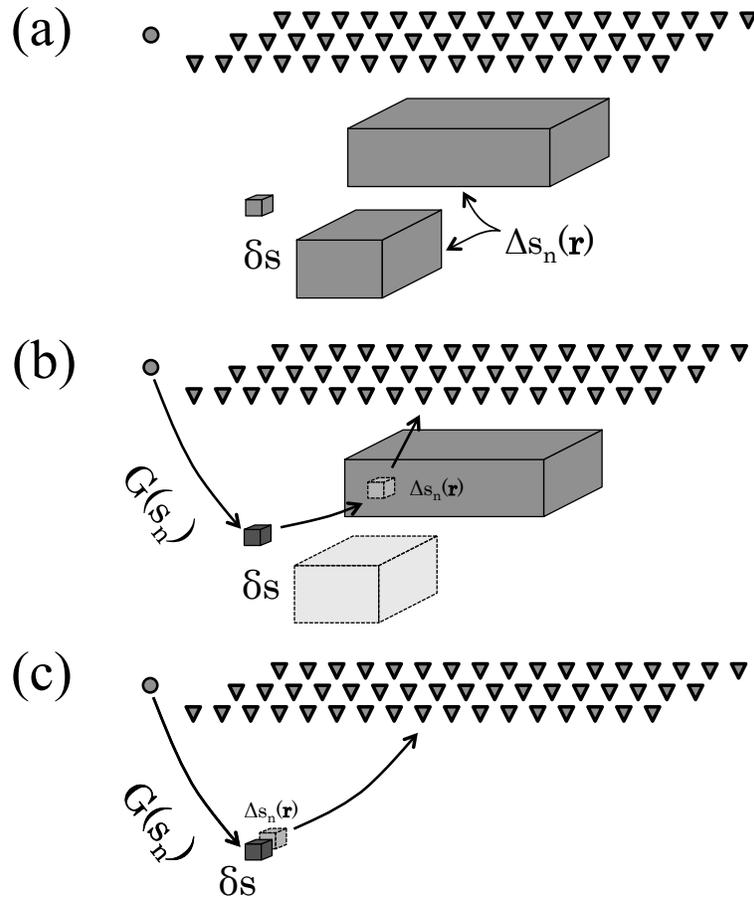
FIG. 5. (a) The new sensitivity calculation is based on the standard variation point $\delta s$, but also the influence of model residuals $\Delta s_n$. The framework allows us to retain or reject a broad range of scattering processes between these quantities to form approximations to full nonlinear sensitivities: (b) for instance we can include only model residuals corresponding to unreconstructed overburden structures, or (c) we can include only model residuals coinciding spatially with the variation point under study.

where $\Delta s_n(\mathbf{r}) = s_{n+1}(\mathbf{r}) - s_n(\mathbf{r})$ is the difference between the $n + 1$th and the $n$th media. To construct the second series, we add to $s_{n+1}$ on the left a variation $\delta s$ localized at the position $\mathbf{r}$, balancing the right side by adding the same quantity to each instance of $\Delta s_n$:

$$
\begin{aligned}
G(\mathbf{r}_g, \mathbf{r}_s | s_{n+1} + \delta s) = {}& G(\mathbf{r}_g, \mathbf{r}_s | s_n) \\
& - \omega^2 \int d\mathbf{r}' G(\mathbf{r}_g, \mathbf{r}' | s_n) \left[ \Delta s_n(\mathbf{r}') + \delta s(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}') \right] G(\mathbf{r}', \mathbf{r}_s | s_n) \\
& + \omega^4 \int d\mathbf{r}' G(\mathbf{r}_g, \mathbf{r}' | s_n) \left[ \Delta s_n(\mathbf{r}') + \delta s(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}') \right] \\
& \times \int d\mathbf{r}'' G(\mathbf{r}', \mathbf{r}'' | s_n) \left[ \Delta s_n(\mathbf{r}'') + \delta s(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}'') \right] G(\mathbf{r}'', \mathbf{r}_s | s_n) + ...
\end{aligned}
\tag{11}
$$

The variation $\delta G(s_{n+1})$ is the difference between these two series:

$$
\delta G(\mathbf{r}_g, \mathbf{r}_s | s_{n+1}) = [G_0' + G_{11}' + G_{12}' + ...] \, \delta s(\mathbf{r}) + ...,
\tag{12}
$$

where

$$
\begin{aligned}
G_0' &= -\omega^2 G(\mathbf{r}_g, \mathbf{r} | s_n) G(\mathbf{r}, \mathbf{r}_s | s_n) \\
G_{11}' &= \omega^4 \int d\mathbf{r}' G(\mathbf{r}_g, \mathbf{r} | s_n) G(\mathbf{r}, \mathbf{r}' | s_n) G(\mathbf{r}', \mathbf{r}_s | s_n) \Delta s_n(\mathbf{r}') \\
G_{12}' &= \omega^4 \int d\mathbf{r}' G(\mathbf{r}_g, \mathbf{r}' | s_n) G(\mathbf{r}', \mathbf{r} | s_n) G(\mathbf{r}, \mathbf{r}_s | s_n) \Delta s_n(\mathbf{r}').
\end{aligned}
\tag{13}
$$

These terms are interpretable in terms of scattering processes as illustrated in Figure 6. $\Delta s_n$ and $\delta s$ both represent deviations from $s_n$, and so both act as scatterers. In the full series for $\delta G(s_{n+1})$, scattering processes, like $G_0'$, $G_{11}'$, and $G_{12}'$, which involve one interaction with $\delta s$ (see Figures 6a-c) and multiple interactions with $\Delta s_n$, contribute to the sensitivities.

**Exchanging $\Delta s_n(\mathbf{r})$ for the $n$th residuals through direct nonlinear inverse scattering**

Inverse scattering theory can be used to develop a series relationship between $\delta P^*(s_n)$ and $\Delta s_n(\mathbf{r})$. The complex conjugate of the residuals can be expressed as

$$
\begin{aligned}
\delta P^*(\mathbf{r}_g, \mathbf{r}_s | s_n) &= \int d\mathbf{r}' G^*(\mathbf{r}_g, \mathbf{r}' | s_n) \Delta s_n(\mathbf{r}') G^*(\mathbf{r}', \mathbf{r}_s | s_n) + ... \\
&= \mathcal{G}^* \Delta s_n(\mathbf{r}) + ...,
\end{aligned}
\tag{14}
$$

where in the second line the integral and Green's functions have been collected into the operator $\mathcal{G}^*$. This series is reverted using standard inverse scattering series techniques (Weglein et al., 2003), producing a series expression for $\Delta s_n$ in orders of the residuals:

$$
\Delta s_n(\mathbf{r}) = \mathcal{G}^{*-1} \delta P^*(s_n) + ... \; .
\tag{15}
$$

Substituting equation (15) for $\Delta s_n$ in equation (13) generates sensitivities of the form:

$$
\left( \frac{\partial G(s_{n+1})}{\partial s(\mathbf{r})} \right) = \left( \frac{\partial G(s_{n+1})}{\partial s(\mathbf{r})} \right)_0 + \left( \frac{\partial G(s_{n+1})}{\partial s(\mathbf{r})} \right)_1 + ...,
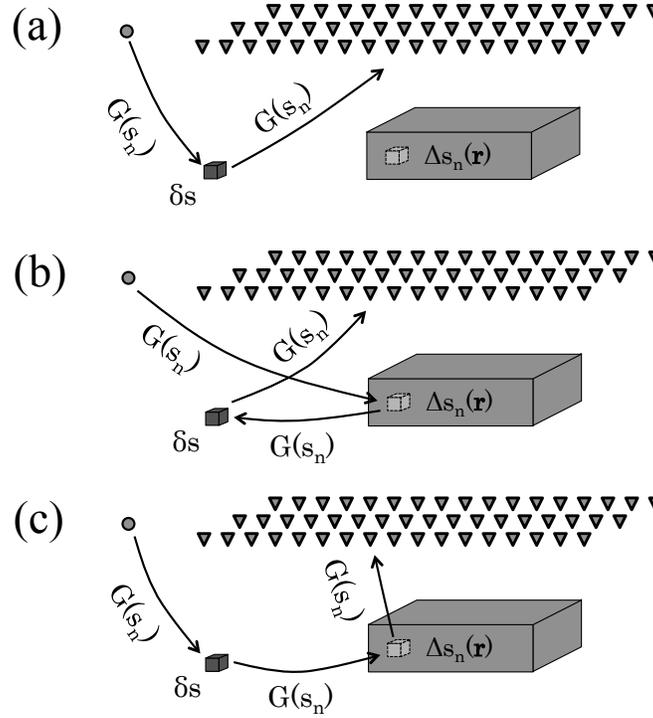\tag{16}
$$

FIG. 6. Scattering processes associated with nonlinear sensitivity calculation: (a) zeroth order in $\Delta s_n$; (b-c) first order in $\Delta s_n$. Propagation via $G(s_n)$ is indicated with a blue arrow.

where the index refers to the order of the term in the residuals $\delta P^*(s_n)$. In the zeroth order term standard FWI sensitivities are recovered:

$$\left(\frac{\partial G(s_{n+1})}{\partial s(\mathbf{r})}\right)_0 = \left(\frac{\partial G(s_n)}{\partial s(\mathbf{r})}\right) = -\omega^2 G(\mathbf{r}_g, \mathbf{r}|s_n)G(\mathbf{r}, \mathbf{r}_s|s_n).$$

The first order term derives from $G'_{11}$ and $G'_{12}$:

$$\begin{aligned}
\left(\frac{\partial G(s_{n+1})}{\partial s(\mathbf{r})}\right)_1 =& \omega^4 \int d\mathbf{r}' \left[ G(\mathbf{r}_g, \mathbf{r}|s_n)G(\mathbf{r}, \mathbf{r}'|s_n)G(\mathbf{r}', \mathbf{r}_s|s_n) \right. \\
& \left. + G(\mathbf{r}_g, \mathbf{r}'|s_n)G(\mathbf{r}', \mathbf{r}|s_n)G(\mathbf{r}, \mathbf{r}_s|s_n) \right] \mathcal{G}^{*-1}\delta P^*(s_n).
\end{aligned} \tag{17}$$

Approximations are arrived at by truncating the series in equation (16).

## CASE I (REVIEW): LARGE CONTRAST/ANGLE REFLECTIVITY

Let us next examine a special case of these nonlinear sensitivities, wherein the resulting gradient (1) is second order in the residuals, and (2) simply incorporates nonlinear reflectivity information. This is achieved by truncating equation (16) at order 1, and considering only a portion of the the full difference $\Delta s_n(\mathbf{r})$ between the $n$th and the $n + 1$th iterate, as illustrated in Figure 5c.

### Second order, collocated scattering approximation

The general scattering picture (see Figures 6a-c) is replaced with a $\Delta s_n(\mathbf{r})$ that is localized and collocated with the variation point. This is obtained by setting $\Delta s_n(\mathbf{r}') =$

$\Delta s_n(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}')$ such that second term in equation (16) becomes

$$\left(\frac{\partial G(s_{n+1})}{\partial s(\mathbf{r})}\right)_1 = 2\omega^4 G(\mathbf{r}_g, \mathbf{r}|s_n)G(\mathbf{r}, \mathbf{r}_s|s_n)G(\mathbf{r}, \mathbf{r}|s_n)\mathcal{G}^{*-1}\delta P^*(s_n),$$

and the quantity $\mathcal{G}^{*-1}\delta P^*(s_n)$ reduces to

$$\mathcal{G}^{*-1}\delta P^*(s_n) = -\frac{1}{\omega^2}\frac{\delta P^*(\mathbf{r}_g, \mathbf{r}_s|s_n)}{G^*(\mathbf{r}_g, \mathbf{r}|s_n)G^*(\mathbf{r}, \mathbf{r}_s|s_n)}.$$

Putting the lowest two orders of the sensitivity together the following case of nonlinear sensitivities is obtained:

$$\left(\frac{\partial G(s_{n+1})}{\partial s(\mathbf{r})}\right) \approx -\omega^2 J_{3D}(\mathbf{r}|s_n)\left[1 + 2\frac{\delta P^*(s_n)}{I_{3D}(\mathbf{r}|s_n)}\right], \tag{18}$$

where

$$I_{3D}(\mathbf{r}|s_n) = \frac{G^*(\mathbf{r}_g, \mathbf{r}|s_n)G^*(\mathbf{r}, \mathbf{r}_s|s_n)}{G(\mathbf{r}, \mathbf{r}|s_n)}, \quad \text{and}$$
$$J_{3D}(\mathbf{r}|s_n) = G(\mathbf{r}_g, \mathbf{r}|s_n)G(\mathbf{r}, \mathbf{r}_s|s_n). \tag{19}$$

$G(\mathbf{r}, \mathbf{r}|s_n)$ is singular, implying that in 3D a principle value for the gradient integral will be required; in the 1.5D cases no poles appear.

## 1.5D form of equation (18)

In order to verify that they meaningfully incorporate nonlinear information, we consider a 1.5D version of the sensitivity formula in equation (18). In 1.5D the medium varies in depth only, and thus sensitivities are defined in terms of medium variations in $z$, but the wave physics is 2D (coordinates $x$ and $z$). Under this restriction, the associated 2nd order sensitivity formula has the same essential form,

$$\left(\frac{\partial G(s_{n+1})}{\partial s(z)}\right) \approx -\omega^2 J_{1D}(z|s_n)\left[1 + 2\frac{\delta P^*(k_g, \omega)}{I_{1D}(z|s_n)}\right], \tag{20}$$

but some slight differences in the weights:

$$J_{1D}(z|s_n) = \int dx' G(k_g, 0, x', z|s_n)G(x', z, 0, 0|s_n),$$

and

$$I_{1D}(z|s_n) = \frac{J_{1D}^*(z|s_n)J_{1D}(z|s_n)}{\int dx' \int dx'' G(k_g, 0, x', z|s_n)G(x', z, x'', z|s_n)G(x'', z, 0, 0|s_n)}.$$

## Nonlinear sensitivities for the first 1.5D FWI update

For the purposes of analysis we will consider FWI updates derived from residuals in the $(k_g, \omega)$ domain (i.e., one shot record of data Fourier transformed over geophone position and time), holding $k_g$ fixed. This will allow us to distinguish between updating with high angle (large $k_g$) vs. low angle data. The objective function is modified to

$$\phi(s_n) = \frac{1}{2} \int d\omega |\delta P(k_g, \omega)|^2, \tag{21}$$

and it is minimized with updates of Gauss-Newton form:

$$\delta s_0(z) = -\int dz' H_0^{-1}(z, z') g_0(z'), \tag{22}$$

where the gradient is based on a version of the nonlinear sensitivities:

$$g_n(z) = -\int d\omega \left( \frac{\partial G(s_{n+1})}{\partial s(z)} \right) \delta P_n^*(k_g, \omega), \tag{23}$$

with $n = 0$; the Hessian is based on standard sensitivities:

$$H_n(z, z') = \int d\omega \frac{\partial G(s_n)}{\partial s(z)} \frac{\partial G^*(s_n)}{\partial s(z')}, \tag{24}$$

also with $n = 0$. If the initial medium is homogeneous, we can analyze this update using exact forms for the Green's functions (Clayton and Stolt, 1981): $G(k_g, z_g, x, z|s_0) = (i2q_g)^{-1} e^{-ik_g x + iq_g|z_g - z|}$ and $G(x, z, x_s, z_s|s_0) = (2\pi)^2 \int dk' (i2q')^{-1} e^{ik'(x-x_s)+iq'|z'-z_s|}$, where $q_\alpha^2 = \omega^2 s_0 - k_\alpha^2$. Adjusting the shot record coordinate system such that $z_g = z_s = x_s = 0$, the $I$ and $J$ reduces

$$J_{1D}(z|s_0) = \frac{e^{i2q_g z}}{(i2q_g)^2}, \;\; I_{1D}(z|s_0) = \frac{e^{-i2q_g z}}{i2q_g}, \tag{25}$$

and the sensitivity itself becomes

$$\left( \frac{\partial G(s_1)}{\partial s(z)} \right) = -\omega^2 \left[ \frac{e^{i2q_g z}}{(i2q_g)^2} \right] \left[ 1 + 2\delta P^*(k_g, \omega)(i2q_g) e^{i2q_g z} \right]. \tag{26}$$

Reconstruction of a single-interface model in is considered. The goal is the determination, from a constant initial medium $s_0$, of the profile $s(z) = s_0 + \Delta s S(z - z_1)$, where $S$ is a step or Heaviside function. $\Delta s$ is the amplitude of the ideal update, taking us directly from the initial model to the correct answer. The backscattered amplitude (i.e., the reflection coefficient) can be expressed as a series in orders of this $\Delta s$:

$$R(\theta) = R_1(\theta) + R_2(\theta) + ..., \;\; \text{where} \tag{27}$$

$$R_1(\theta) = -\frac{1}{4} \left( 1 + \sin^2 \theta \right) \left( \frac{\Delta s}{s_0} \right), \;\; R_2(\theta) = \frac{1}{8} \left( 1 + 2\sin^2 \theta \right) \left( \frac{\Delta s}{s_0} \right)^2,$$
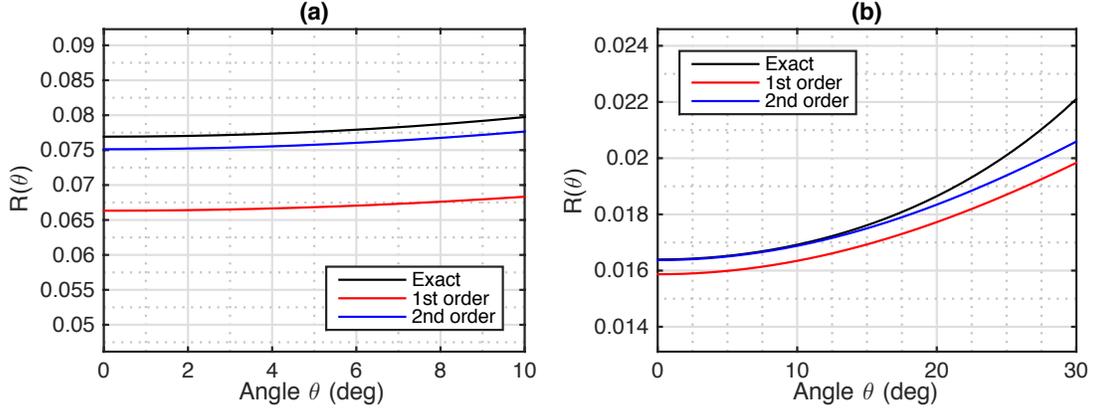
FIG. 7. (a) $R(\theta)$ for low angles but large contrasts; (b) $R(\theta)$ for small contrasts but large angles.

etc. In Figure 7a, we note by comparing exact, first order ($R \approx R_1$) and second order ($R \approx R_1 + R_2$) coefficient calculations that basic linearization error can arise at low angles when contrasts are large, and (in Figure 7b) at high angle even when contrasts are low. We also note that corrections out to as low as second order can lead to significant error reduction, however. With analyzable formulas for $R$ in hand, we can then analytically express the complex conjugate of the $(k_g, \omega)$ domain residuals at the first iteration (Innanen, 2014a):

$$\delta P_0^*(k_g, \omega) = -R(\theta)\frac{e^{-i2q_g z_1}}{i2q_g}. \tag{28}$$

**Response of 2nd-order sensitivities to backscattered data**

With all the ingredients for the sensitivities now available in analytic form, we may analyze the first iteration in the reconstruction of Figure **??**a. The gradient now has two terms, one 1st order in $\delta P^*$ and the other 2nd, that is, $g_0(z) = g_0^{(1)}(z) + g_0^{(1)}(z)$, where

$$
\begin{aligned}
g_0^{(1)}(z) &= \frac{c_0^2}{4}R(\theta) \int d\omega \left(\frac{\omega^2/c_0^2}{q_g^2}\right) \frac{e^{i2q_g(z-z_1)}}{i2q_g}, \text{ and} \\
g_0^{(2)}(z) &= -\frac{c_0^2}{2}R^2(\theta) \int d\omega \left(\frac{\omega^2/c_0^2}{q_g^2}\right) \frac{e^{i2q_g(2z-2z_1)}}{i2q_g}.
\end{aligned} \tag{29}
$$

Noting that (Innanen, 2014a) $d\omega = d2q_g(c_0/2\cos\theta)$ and $q_g = (\omega/c_0)\cos\theta$, we can evaluate these integrals and reassemble the gradient, obtaining

$$g_0(z) = \frac{c_0^3\pi}{4\cos^3\theta}[R(\theta) - 2R^2(\theta)]S(z - z_1). \tag{30}$$

The Hessian, which we have given a standard Gauss-Newton approximate form (e.g., Virieux and Operto, 2009), evaluates in this simple environment (Innanen, 2014a) to

$$H_0(z, z') = c_0^5\pi(16\cos^5\theta)^{-1}\delta(z - z'), \tag{31}$$

and so, via equation (22), the update is of the form

$$\delta s_0(z) = -\frac{4\cos^2\theta}{c_0^2}\left[R(\theta) - 2R^2(\theta)\right]S(z - z_1).\tag{32}$$

### Comparison of 2nd order vs. standard Gauss-Newton update

We characterized the ideal update as $\Delta s(z) = \Delta s S(z - z_1)$ and related it to the reflection coefficient through equation (15). To analyze the relative accuracy of the first order and second order FWI iterations, we will substitute two truncations of the series for $R(\theta)$ into equation (32). The standard Gauss-Newton update is recovered by neglecting $R^2$; noting also that to leading order in $\sin^2\theta$ we may replace $1/\cos^2\theta$ with $(1 + \sin^2\theta)$, we obtain

$$\left(\frac{\delta s_0(z)}{s_0}\right)_1 \approx -\left(\frac{4}{1 + \sin^2\theta}\right)R(\theta)S(z - z_1).\tag{33}$$

The nonlinear Gauss-Newton-like update, based on second order collocated sensitivities, is

$$\left(\frac{\delta s_0(z)}{s_0}\right)_2 = -\left(\frac{4}{1 + \sin^2\theta}\right)\left[R(\theta) - 2R^2(\theta)\right]S(z - z_1).\tag{34}$$

Let us first verify that a standard linear Gauss-Newton update is equivalent to the ideal update to 1st order. If contrasts and angles are low, 2nd order contributions to $R$ are neligible, and the reflection coefficient is

$$R(\theta) \approx R_1(\theta) = -\frac{1}{4}\left(1 + \sin^2\theta\right)\left(\frac{\Delta s}{s_0}\right);\tag{35}$$

substituting this into equation (33) we obtain

$$\left(\frac{\delta s_0(z)}{s_0}\right)_1 \approx \left(\frac{\Delta s}{s_0}\right)S(z - z_1),\tag{36}$$

demonstrating the equivalence of $\delta s_0(z)$ and the ideal $\Delta s(z)$. However, if the angle or contrast is such that second order contributions to $R(\theta)$ are non-negligible, referring to equation (15) we must instead substitute

$$R(\theta) \approx -\frac{1}{4}\left(1 + \sin^2\theta\right)\left(\frac{\Delta s}{s_0}\right) + \frac{1}{8}\left(1 + 2\sin^2\theta\right)\left(\frac{\Delta s}{s_0}\right)^2,\tag{37}$$

and this produces a discrepancy at second order between the Gauss-Newton update $\delta s_0(z)$ and the ideal update $\Delta s(z)$:

$$\left(\frac{\delta s_0(z)}{s_0}\right)_1 \approx \left[\left(\frac{\Delta s}{s_0}\right) - \frac{1}{2}(1 + \sin^2\theta)\left(\frac{\Delta s}{s_0}\right)^2\right]S(z - z_1).\tag{38}$$

Let us compare this with the update generated using the second order sensitivity expression. Substituting the reflection coefficient in equation (37) into equation (34), a corrective term
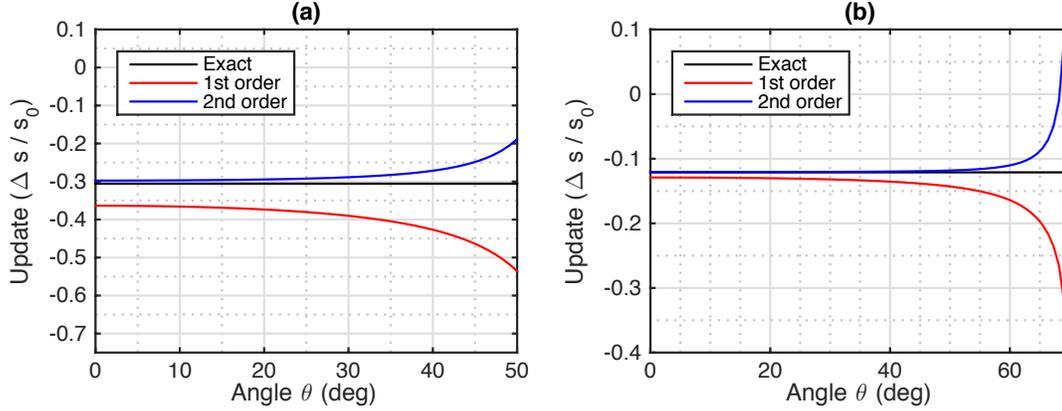
FIG. 8. Exact (black), first order (red), vs. second order (blue) updates at each angle of incidence.

is introduced at second order, exactly suppressing the second order discrepancy corrupting equation (38), such that the resulting update lapses back to

$$\left(\frac{\delta s_0(z)}{s_0}\right)_2 \approx \left(\frac{\Delta s}{s_0}\right) S(z - z_1); \tag{39}$$

here the consistency of the candidate and ideal updates extends to second order, rather than just first. In Figure 8 the difference between 2nd order sensitivities and (standard) 1st order sensitivities is illustrated. Because we consider a fixed $k_g$, we can examine the accuracy angle by angle; a full inversion would sum over $k_g$ and thus average over these angles. In Figure 8a the interface is large contrast, going from $c_0 = (s_0)^{-1/2} = 1500$m/s in the upper halfspace to $c_1 = (s_1)^{-1/2} = 1800$m/s in the lower; especially in the range $\theta = 0°$-$30°$ the difference between the standard Gauss-Newton update and that based on second order sensitivities is significant. Meanwhile in Figure 8b the interface represents a small contrast, with $c_0 = 1500$m/s, $c_1 = 1600$m/s, but is examined over a wider range of angles. Here the second order update "sticks to" the exact update to roughly $\theta = 60°$, in contrast to the standard update which deviates significantly at $\theta = 30°$.

## CASE II: LARGE/EXTENDED MODEL RESIDUALS

In Case I, we took the general sensitivity in equation (16) and specified it in three ways: the dimensionality was reduced to 1.5D, the maximum order of scattering was limited to two, and the scattering between the variation point ($\delta s$) and the portions of the model being constructed ($\Delta s$ and/or $\delta P$) was restricted to that for which the two were collocated in space. This is seen to suffice for problems involving nonlinear treatment of reflection amplitudes. In this second situation, we relax the restrictions somewhat (see Figure 5b). Now, scattering will be permitted to occur between variation points and model update points separated in space. Furthermore, an infinite number of these interactions will be included, a subset of the entire range of possible scattering processes.

**Forming a gradient sensitive to the unreconstructed overburden**

We will restrict the discussion to a fully 1D environment, in which the sensitivity expression has the general form

$$\left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right) = \left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_0 + \left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_1 + \left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_2 + ..., \quad (40)$$

with the zero'th order term being the standard

$$\left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_0 = -\omega^2 G_n(z_g, z) G_n(z, z_s). \quad (41)$$

The general first and second order terms are then

$$\left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_1 = \omega^4 \left[\int dz' G_n(z_g, z') \Delta s_n(z') G_n(z', z)\right] G_n(z, z_s)$$
$$+ \omega^4 G_n(z_g, z) \left[\int dz' G_n(z, z') \Delta s_n(z') G_n(z', z_s)\right], \quad (42)$$

at first order, and

$$\left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_2 = \left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_2^1 + \left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_2^2 + \left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_2^3, \quad (43)$$

at second order, where

$$\left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_2^1 = -\omega^6 \left[\int dz' G_n(z_g, z') \Delta s_n(z') \int dz'' G_n(z', z'') \Delta s_n(z'') G_n(z'', z)\right] G_n(z, z_s),$$

and

$$\left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_2^2 = -\omega^6 \left[\int dz' G_n(z_g, z') \Delta s_n(z') G_n(z', z)\right] \left[\int dz'' G_n(z, z'') \Delta s_n(z'') G_n(z'', z_s)\right],$$

and

$$\left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_2^1 = -\omega^6 G(z_g, z) \left[\int dz' G_n(z, z') \Delta s_n(z') \int dz'' G_n(z', z'') \Delta s_n(z'') G_n(z'', z_s)\right].$$

The flexibility of this approach lies in our ability to associate certain scattering geometries with each of these terms. The terms as they stand in equations (42) and (43) are divided up based on where, in the 2 or 3 scattering interactions, the variational perturbation in the model occurs, relative to the update perturbation. See Figure 10.

We do not want to include the entire range of nonlinearity within the sensitivity, but, instead, investigate it for its ability to allow us to include this or that feature of the full nonlinearity. Let us use, for the sake of experimentation, only those terms corresponding with the sensitivity of the update to unreconstructed model variations in the overburden,
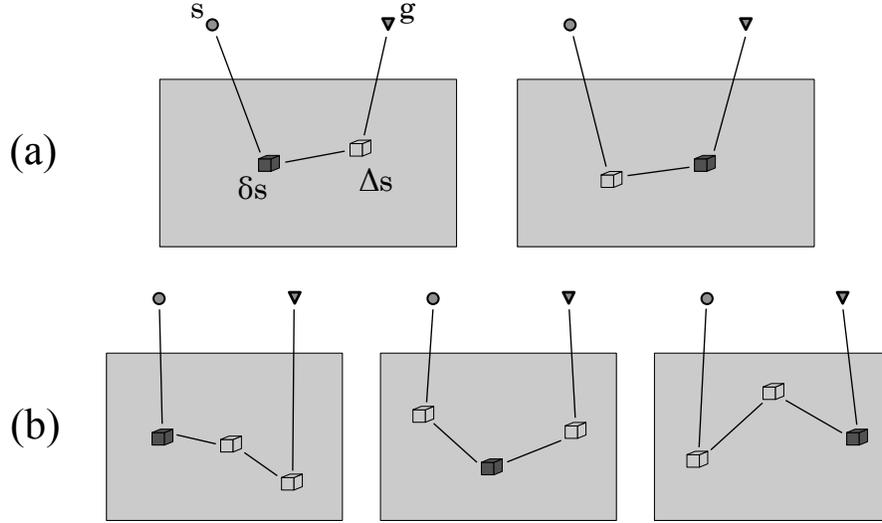
FIG. 9. Scattering processes involved in the first and second order 1D sensitivity calculation. (a) Left to right respectively, the first and second sensitivity components in equation (42) respectively. (b) Left to right, components (1), (2) and (3) in equation (43).

that is, in model regions above the current variation point. Mathematically, this means including scattering between $\delta s(z)$ and $\Delta s_0(z')$ when $z > z'$, but rejecting contributions when $z' > z$.

To make a straightforwardly analyzable example, let us consider the problem of constructing the first FWI update on a homogeneous background. Using the Green's function $G(z_g, z_s) = (i2\omega/c_0)^{-1} \exp i2\omega|z_g - z_s|/c_0$, the zeroth order term evaluates to

$$\left( \frac{\partial G(z_g, z_s)}{\partial s(z)} \right)_0 = \frac{c_0}{4} e^{i2kz}, \tag{44}$$

and the first order term likewise evaluates to

$$\left( \frac{\partial G(z_g, z_s)}{\partial s(z)} \right)_1 = \frac{i}{4} \omega c_0^3 e^{i2kz} \int_{-\infty}^{z} dz' \Delta s_0(z')$$
$$+ \frac{i}{4} \omega c_0^3 \int_{z}^{\infty} dz' e^{i2kz'} \Delta s_0(z'). \tag{45}$$

The two terms in equation (45) can be interpreted in the lower-vs-higher terms required to apply the accept/reject strategy above. Inspection of the the integrals indicates that the first term counts up contributions from the $\Delta s_0(z')$ perturbation at depths shallower than the output point $z$, and the second term counts up contributions from $\Delta s_0(z')$ values below the output point $z$. The output point is the point at which the sensitivity is being evaluated. Thus, the first term has the scattering diagram illustrated in Figure 10a, and the second has the scattering diagram illustrated in Figure 10b.

Third order terms can be evaluated and analyzed similarly. In fact, comparing the zeroth, first, and second order terms fulfilling our condition that only $\Delta s_0$ contributions from *above* the variation point are accepted (i.e., Figure 10a but not b), a mathematical
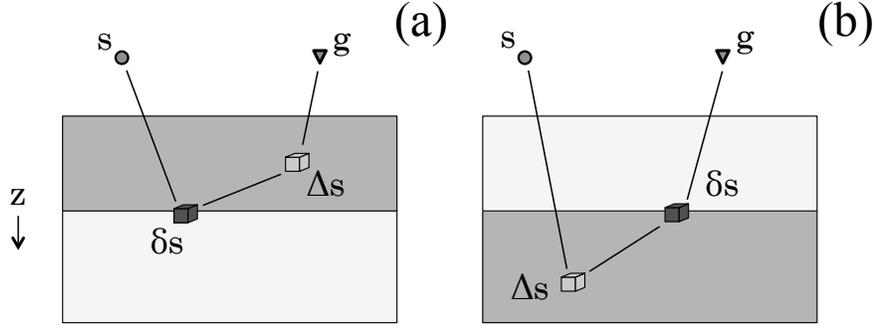
FIG. 10. Ordering in depth of second order scattering. (a) Sensitivity to structures above the current variation point, i.e., the as yet unreconstructed overburden (dark grey region). (b) Sensitivity to structures below the current variation point, i.e., the as yet unreconstructed underburden (dark grey region).

pattern appears whose form is predictable for the $n$th order term. Selecting only these terms would amount to a calculation of the sensitivity to the as-yet unreconstructed overburden. Retaining all terms obeying this restriction on scattering geometry, we obtain an overburden sensitivity we will label with a $+$ sign:

$$\left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_+ = \frac{c_0^2}{4}e^{i2kz}\left\{1 + i\omega c_0 \int_0^z dz'\Delta s_0(z') - \omega^2 c_0^2 \frac{1}{2}\left[\int_0^z dz'\Delta s_0(z')\right]^2 + ...\right\},$$

or, more simply,

$$\left(\frac{\partial G(z_g, z_s)}{\partial s(z)}\right)_+ = \frac{1}{4s_0}\exp\left\{i\omega s_0^{1/2}\left[z + \frac{1}{2}\int_0^z dz'\frac{\Delta s_0(z')}{s_0}\right]\right\}. \tag{46}$$

In the last line we have recognized the sensitivity calculation as the Taylor's series for a complex exponential, and replaced all velocities with the counterpart model parameter $s_0$. So, in this simple case study, a closed-form is obtainable, and thus we can consider the implicit effect of an infinite number of contributions of $\Delta s_0(z')$, all overlying the variation point, to the sensitivity.

### Substitution of $\Delta s_0(z)$ with the residuals using direct nonlinear inverse scattering

As we discussed in the general approach, the result is interesting but not yet of direct value, because during the first iteration we do not yet know what $\Delta s_0(z)$ is. We will view it as an opportunity to involve the residuals, and techniques of direction nonlinear inverse scattering simultaneously. For instance, Innanen (2008) showed that a large, extended overburden perturbation can be related through a straightforwardly computable series to the scattered field. The relationship between the residuals and the update under construction is almost identical to this series, and so we may write:

$$\left(\frac{\Delta s(z)}{s_0}\right) = \frac{1}{0!}\left\{4\delta P_0(z)\right\} - \frac{1}{1!}\left\{8\frac{d}{dz}\left[\delta P_0(z)\int_0^z dz'\delta P_0(z')\right]\right\}$$
$$+ \frac{1}{2!}\left\{16\frac{d^2}{dz^2}\left[\delta P_0(z)\left(\int_0^z dz'\delta P_0(z')\right)^2\right]\right\} - ..., \tag{47}$$

where $z = c_0 t/2$ is a pseudodepth, and we have constructed $\delta P(z)$ by inverse Fourier transforming $\delta P(\omega)$ over frequency and changed variables from time $t$ to $z$ using the pseudodepth definition. We may then use this formula to replace $\Delta s_0(z)$ in the nonlinear sensitivity with, in principle, any desired order of the series in equation (46). To leading order, for instance, we produce

$$\left( \frac{\partial G(z_g, z_s)}{\partial s(z)} \right)_+ = \frac{c_0^2}{4} \exp \left\{ i2\omega s_0^{1/2} \left[ z + 2 \int_0^z dz' \delta P_0(z') \right] \right\}. \tag{48}$$

Notice that in the limit $\delta P_0(z) \to 0$, this sensitivity lapses to the standard FWI sensitivity form (see, e.g., equation 44)

$$\lim_{\delta P \to 0} \left( \frac{\partial G(z_g, z_s)}{\partial s(z)} \right)_+ = \frac{c_0^2}{4} \exp \left( i2\omega s_0^{1/2} z \right). \tag{49}$$

**Nonlinear gradient**

This means our extended gradient becomes

$$\begin{aligned}
g_0(z) &= -\int d\omega \left( \frac{\partial G(z_g, z_s)}{\partial s(z)} \right)_+ \delta P_0^*(\omega) \\
&= \frac{c_0^2}{4} \int d\omega \exp \left\{ i2\omega s_0^{1/2} \left[ z + 2 \int_0^z dz' \delta P_0(z') \right] \right\} \delta P_0^*(\omega).
\end{aligned} \tag{50}$$

The change imparted to the normal, linear sensitivity in this new extended form is primarily in the locations of the interfaces in the update, as dictated by the residuals, which will be altered with the integral term in the argument of the exponential. It can be shown that the lower interfaces, which are placed at depths consistent with wave propagation at the initial medium velocity, are incorrect by a factor of $c_{\text{ave}}/c_0 \times$overburden thickness. Because at depths greater than $z_1$,

$$2 \int_0^z dz' \delta P_0(z') = 2R_1(z - z_1), \tag{51}$$

the placement of a deeper interface $z_2$ is altered by a shift proportional to $2R_1 \times$ the contributing depth below $z_1$. Now, since the error ratio $c_{\text{ave}}/c_0$ is equal to $(1 + 2R_1)$ to leading order in $R_1$, the effect is to place the deeper reflector in such a way that the error associated with overlying model residuals is countered to a high degree of precision. So, the result of extending the FWI sensitivities to accommodate the overburden *we are in the process of constructing*, leads to a first update that has brought us much closer to the final result than the standard update would have.

## CONCLUSIONS

Nonlinear, residual dependent FWI sensitivities are the outgrowth of the observation that certain direct nonlinear procedures, available in special case environments/schemes, such as AVO inversion or 1D/1.5D direct nonlinear inverse scattering imaging and inversion, have no FWI generalization – they cannot be found as special cases of any standard

FWI procedure. In order to derive general FWI schemes that reduce in this way, an extension of some kind is required to our ideas of FWI sensitivity. A proposal from 2014, namely that the sensitivities be created by varying not the current model iteration, but the forthcoming iteration, is pursued in this paper to confirm that this approach correctly reduces to an existing 1D normal incidence direct nonlinear imaging and inversion scheme, previously derived from inverse scattering considerations.

## ACKNOWLEDGEMENTS

## REFERENCES

Abramowitz, M., and Stegun, I. A., 1972, Handbook of mathematical functions: Dover, 9th edn.

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

Innanen, K. A., 2008, A direct non-linear inversion of primary wave data reflecting from extended, heterogeneous media: Inverse Problems, , No. 24, 035,021.

Innanen, K. A., 2009, Born series forward modeling of seismic primary and multiple reflections: an inverse scattering shortcut: Geophys. J. Int., **177**, No. 3, 1197–1204.

Innanen, K. A., 2011, Inversion of the seismic AVF/AVA signatures of highly attenuative targets: Geophysics, **76**, No. 1, R1–R11.

Innanen, K. A., 2014a, Seismic AVO and the inverse Hessian in precritical reflection full waveform inversion: Geophysical Journal International, **199**, 717–734.

Innanen, K. A., 2014b, Seismic full waveform inversion with nonlinear sensitivities: CREWES Annual Report, **26**.

Malcolm, A., and de Hoop, M. V., 2005, A method for inverse scattering based on the generalized bremmer coupling series: Inverse Problems, **21**, 1137–1167.

Shaw, S. A., Weglein, A. B., Foster, D. J., Matson, K. H., and Keys, R. G., 2004, Isolation of a leading order depth imaging series and analysis of its convergence properties: Journal of Seismic Exploration, **2**, 157–195.

Stolt, R. H., and Jacobs, B., 1980, An approach to the inverse seismic problem: Stanford Exploration Project, **25**.

Virieux, J., and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, No. 6, WCC1.

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, , No. 19, R27–R83.

Weglein, A. B., Boyce, W. E., and Anderson, J. E., 1981, Obtaining three-dimensional velocity information directly from reflection seismic data: An inverse scattering formalism: Geophysics, **46**, No. 8, 1116–1120.

Weglein, A. B., Gasparotto, F. A., Carvalho, P. M., and Stolt, R. H., 1997, An inverse-scattering series method for attenuating multiples in seismic reflection data: Geophysics, **62**, No. 6, 1975–1989.

Zhang, H., 2006, Direct non-linear acoustic and elastic inversion: towards fundamentally new comprehensive and realistic target identification: Ph.D. thesis, University of Houston.

Zhang, H., and Weglein, A. B., 2009a, Direct nonlinear inversion of multiparameter 1D acoustic media using inverse scattering subseries: Geophysics, **74**, No. 6, WCD29–WCD39.

Zhang, H., and Weglein, A. B., 2009b, Direct nonlinear inversion of multiparameter 1D elastic media using the inverse scattering series: Geophysics, **74**, No. 6, WCD15–WCD27.