Toward full waveform inversion for source location and velocity model: gradient and Hessian

Nadine Igonin, Kristopher H. Innanen

ABSTRACT

Full waveform inversion (FWI) is used to reconstruct physical parameters of the subsurface. In this report, we move towards a modified FWI scheme that iteratively solves for both source distribution uncertainty and velocity model uncertainty in the acoustic environment. The sensitivities are derived for these two parameters in both time and frequency domain. Additionally, to begin exploring the cross-talk between these two parameters, the Hessian is derived in the frequency domain. The paper concludes with an analytic example demonstrating the behaviour of this formulation and a discussion on the implications of these results.

INTRODUCTION

Full waveform inversion (FWI) has been demonstrated to be a valuable tool for creating high-resolution images of the subsurface (Tarantola (1984),Virieux and Operto (2009)). The conventional FWI problem aims to resolve certain parameters, such as P- and/or S-wave velocity, density, and more recently, anisotropy (Alkhalifa and Plessix (2014),Pan et al. (2016)). In these applications, it is assumed that the source location and character is known, with the focus being on the medium and not the source or receivers. Herein, we propose to instead consider an application of FWI where the source distribution is also an unknown in the formulation.

One application where this sort of problem is a reality is with microseismic data. Microseismic data refers to earthquakes of magnitude less than zero. These earthquakes are commonly studied in the context of geothermal applications, hydraulic fracturing and mining. Often, the velocity model is poorly constrained, and is the subject of refinement using various methods such as calibration with perforation shots or seismic data (Bardainne and Gaucher, 2010). In practice, simplified models are used, with conventional processors often limiting themselves to 1D blocked models. Although anisotropy and 3D models are more recent advances (Maxwell et al., 2010), obtaining accurate velocity models is a challenge.

Another challenge with microseismic data is the accuracy of event locations. Though dependent on the recording geometry, microseismic events tend to have low signal-to-noise ratio. For this reason, many increasingly complex methods have been developed to locate microseismic events with greater accuracy (Waldhauser and Ellsworth (2000), Artman et al. (2010)). However, there is always a certain, perhaps significant, uncertainty in the location.

Consider a receiver array on the surface, with microseismic events randomly distributed in the subsurface. This is a simple transmission problem, and the arrival times of the microseismic events, as well as their coda, contain information about the subsurface the waves travelled through. Therefore, one can imagine conceiving a FWI formulation where the microseismic data can be used to refine a velocity model (Figure 1). The transmission problem has been studied in detail by Long et al. (2009),Kamath and Tsvankin (2014) and
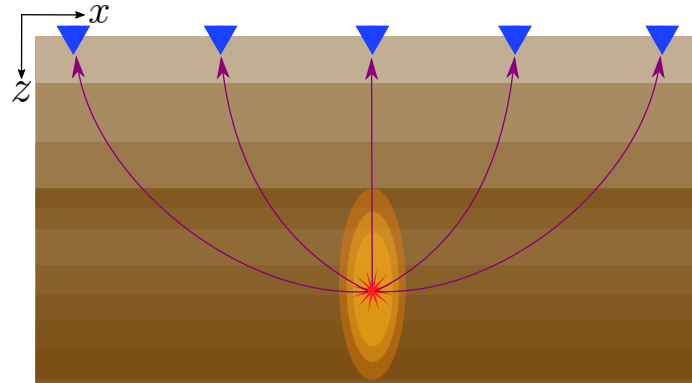
FIG. 1. Schematic illustrating the geometry of a microseismic event (red star) propagating in a layered medium to a set of receivers (blue triangles). The orange ellipses illustrate uncertainty in location, which is especially pronounced in depth for a surface array.

Butzer et al. (2013) in the visco-acoustic, 2D VTI and 3D elastic environment respectively.

Extending the idea further, a similar FWI scheme could be developed where instead of solving for the velocity model, it is the source location and character that is converged upon. Since both the velocity model uncertainty and source distribution uncertainty problem are similar in their wave nature, one could construct a scheme to iteratively solve for both. This would require the construction of a new gradient for the source distribution uncertainty and there would be inherent cross-talk between the two parameters.

The idea of applying FWI to microseismic has become a recent research topic of interest in the international community, but the solution is far from found. Jarillo and Tsvankin (2017) and Sethi and Shekar (2017) have made strides in applying FWI to this problem, though limited to simplistic synthetic examples that want for a more realistic appraisal of microseismic. This is an emerging area of research with much work still to be done.

In this paper, we derive the sensitivities for the source distribution and velocity model uncertainties in both time and frequency domain. Then, to study the effect of cross-talk, we construct the Hessian in frequency domain. To demonstrate how the two sensitivities effect each other, we undertake a simple analytic example. Finally, we conclude with some future work.

THEORY

In the following sections, equations for the gradient and Hessian are derived for the purpose of updating the velocity model and source location iteratively. We start with the frequency domain sensitivities, and in an analogous way, repeat the analysis in the time domain. Finally, we construct the Hessian in the frequency domain and discuss its characteristics.

Frequency domain gradients

For simplicity, we begin this analysis with an acoustic source and restrict ourselves to a constant density model. In the frequency domain case, we assume that we know the origin
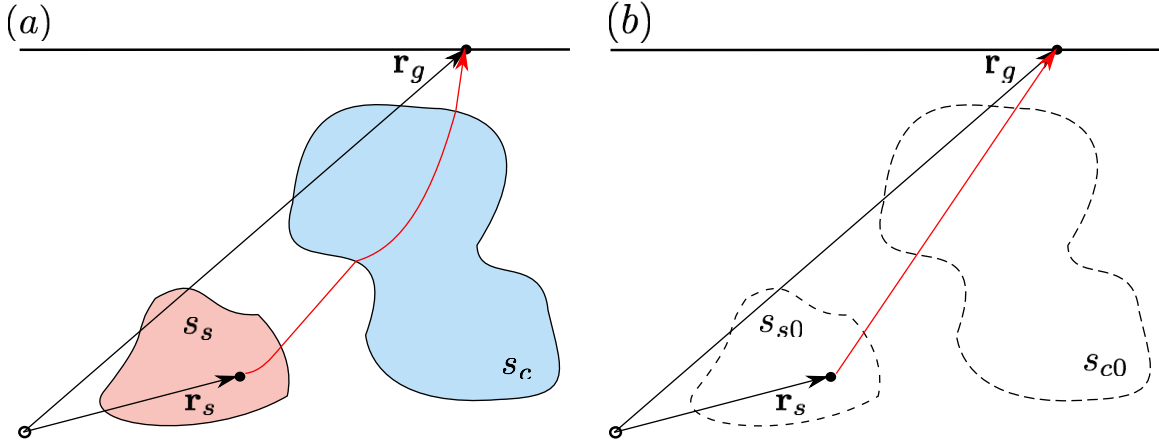
FIG. 2. Schematic illustration of coordinate system and key variables. (a) Red region shows source distribution, while blue region shows the true velocity model distribution. Red line indicates ray path from source to receiver; (b) The same model space where the dashed regions are the background source and velocity distributions.

time of the source, and set it to $t = 0$. The source, $s_s(\mathbf{r})$, has an unknown spatial distribution and emits waves that travel through an unknown scalar acoustic medium $s_c(\mathbf{r}) = \frac{1}{\mathbf{c}^2(\mathbf{r})}$. This wavefield is recorded at several receiver locations $\mathbf{r}_g$. Figure 2a shows a schematic of this scenario.

Given this, we have the acoustic Helmholtz equation in $s_c(\mathbf{r})$ due to $s_s(\mathbf{r})$:

$$\left[\nabla^2 + \omega^2 s_c(\mathbf{r_g})\right] P(\mathbf{r}_g, \omega | s_c, s_s) = s_s(\mathbf{r_g}), \tag{1}$$

where $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ and $P(\mathbf{r}_g, \omega | s_c, s_s)$ is the recorded wavefield. Next consider that the actual source/medium arrangement is a perturbation on a background arrangement $s_{c0}(\mathbf{r})$, $s_{s0}(\mathbf{r})$, as in Figure 2b. In this case, the Helmholtz equation is

$$\left[\nabla^2 + \omega^2 s_{c0}(\mathbf{r}_g)\right] P(\mathbf{r}_g, \omega | s_{c0}, s_{s0}) = s_{s0}(\mathbf{r}_g). \tag{2}$$

Consider further that we have a delta function source, so the above equation simplifies to

$$\left[\nabla^2 + \omega^2 s_{c0}(\mathbf{r}_g)\right] G(\mathbf{r}_g, \mathbf{r}_s, \omega | s_{c0}, s_{s0}) = \delta(\mathbf{r}_g - \mathbf{r}_s), \tag{3}$$

where we now have the Green's function as the wavefield. The way we connect the above two equations is that the field P due to background source s_{s0} and background medium s_{c0} is the *superposition* of the responses from each point within the source distribution:

$$P(\mathbf{r}_g, \omega | s_{c0}, s_{s0}) = \int d\mathbf{r}_s G(\mathbf{r}_g, \mathbf{r}_s, \omega | s_{c0}, s_{s0}) s_{s0}(\mathbf{r}_s). \tag{4}$$

The same can be said for the true source and medium:

$$P(\mathbf{r}_g, \omega | s_c, s_s) = \int d\mathbf{r}_s G(\mathbf{r}_g, \mathbf{r}_s, \omega | s_c, s_s) s_s(\mathbf{r}_s). \tag{5}$$

Let $s_c(\mathbf{r}) = s_{c0}(\mathbf{r}) + \delta s_c(\mathbf{r})$ and $s_s(\mathbf{r}) = s_{s0}(\mathbf{r}) + \delta s_s(\mathbf{r})$. Then, if $\delta P = P(\mathbf{r}_g, \omega | s_c, s_s) - P(\mathbf{r}_g, \omega | s_{c0}, s_{s0})$, we have

$$\delta P(\mathbf{r}_g, \omega | s_{c0}, s_{s0}) = \int d\mathbf{r}_s G(\mathbf{r}_g, \mathbf{r}_s, \omega | s_{c0} + \delta s_c, s_{s0} + \delta s_s)(s_{s0}(\mathbf{r}_s) + \delta s_s(\mathbf{r}_s)) \quad (6)$$

$$- \int d\mathbf{r}_s G(\mathbf{r}_g, \mathbf{r}_s, \omega | s_{c0}, s_{s0}) s_{s0}(\mathbf{r}_s). \quad (7)$$

The Green's function in equation 6 can be expanded using the Born series (Clayton and Stolt, 1981). Continuing from Eq. 7,

$$\delta P(\mathbf{r}_g, \omega | s_{c0}, s_{s0}) = \int d\mathbf{r}_s \Big[G(\mathbf{r}_g, \mathbf{r}_s, \omega | s_{c0}, s_{s0}) \quad (8)$$

$$- \omega^2 \int d\mathbf{r}' G(\mathbf{r}_g, \mathbf{r}', \omega | s_{c0}, s_{s0}) \delta s_c(\mathbf{r}') G(\mathbf{r}', \mathbf{r}_s, \omega | s_{c0}, s_{s0}) \Big] (s_{s0} + \delta s_s) \quad (9)$$

$$- \int d\mathbf{r}_s G(\mathbf{r}_g, \mathbf{r}_s, \omega | s_{c0}, s_{s0}) s_{s0}(\mathbf{r}_s). \quad (10)$$

Multiplying through, we get five terms in total, but the first term cancels with the last term, leaving three terms:

$$= \int d\mathbf{r}_s G(\mathbf{r}_g, \mathbf{r}_s, \omega | s_{co}, s_{s0}) \delta s_s(\mathbf{r}_s) \quad (11)$$

$$-\omega^2 \int d\mathbf{r}_s \left[\int d\mathbf{r}' G(\mathbf{r}_g, \mathbf{r}', \omega | s_{c0}, s_{s0}) \delta s_c(\mathbf{r}') G(\mathbf{r}', \mathbf{r}_s, \omega | s_{c0}, s_{s0}) \right] s_{s0}(\mathbf{r}_s) \quad (12)$$

$$-\omega^2 \int d\mathbf{r}_s \left[\int d\mathbf{r}' G(\mathbf{r}_g, \mathbf{r}', \omega | s_{c0}, s_{s0}) \delta s_c(\mathbf{r}') G(\mathbf{r}', \mathbf{r}_s, \omega | s_{c0}, s_{s0}) \right] \delta s_s(\mathbf{r}_s) \quad (13)$$

Now we are at a point where we can create our sensitivities one at a time by setting $\delta s_s(\mathbf{r}) = 0$ and then considering a local version of the other perturbation, for example $\delta s_s(\mathbf{r}') = \delta s_s(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}')$. Likewise, then set $\delta s_c(\mathbf{r})$ to zero and use the local version of the other perturbation. Let's start by setting $\delta s_c(\mathbf{r}_s) = 0$ and $\delta s_s(\mathbf{r}_s) = \delta s_s(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}')$, which gives

$$\delta P(\mathbf{r}_g, \omega | s_{c0}, s_{s0}) = \int d\mathbf{r}_s G(\mathbf{r}_g, \mathbf{r}_s, \omega | s_{c0}, s_{s0}) \delta s_s(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}_s) \quad (14)$$

$$= G(\mathbf{r}_g, \mathbf{r}_s, \omega | s_{c0}, s_{s0}) \delta s_s(\mathbf{r}), \quad (15)$$

which is technically a truncated series to first order. Getting this into the form of a sensitivity is a matter of rearranging to get

$$\frac{\partial P(\mathbf{r}_g, \omega)}{\partial s_s(\mathbf{r})} = G(\mathbf{r}_g, \mathbf{r}, \omega | s_{c0}, s_{s0}), \quad (16)$$

to first order. This is simply the Green's function, which conceptually makes sense because it accounts for the direct path from the source to the receiver. Schematically, this is illustrated in Figure 3a.

Similarly, the same thing can be done for the second sensitivity by setting $\delta s_s(\mathbf{r}_s) = 0$ and $\delta s_c(\mathbf{r}_s) = \delta s_c(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}')$, which gives:

$$\delta P(\mathbf{r}_g, \omega | s_{c0}, s_{s0}) = -\omega^2 \int d\mathbf{r}_s \int d\mathbf{r}' G(\mathbf{r}_g, \mathbf{r}', \omega | s_{c0}, s_{s0}) \delta s_c(\mathbf{r}') \delta(\mathbf{r} - \mathbf{r}') \cdot \quad (17)$$

$$G(\mathbf{r}', \mathbf{r}_s, \omega | s_{c0}, s_{s0}) s_{s0}(\mathbf{r}_s) \quad (18)$$

$$= \omega^2 \delta s_c(\mathbf{r}) G(\mathbf{r}_g, \mathbf{r}, \omega | s_{c0}, s_{s0}) \int d\mathbf{r}_s G(\mathbf{r}, \mathbf{r}_s, \omega | s_{c0}, s_{s0}) s_{s0}(\mathbf{r}_s). \quad (19)$$
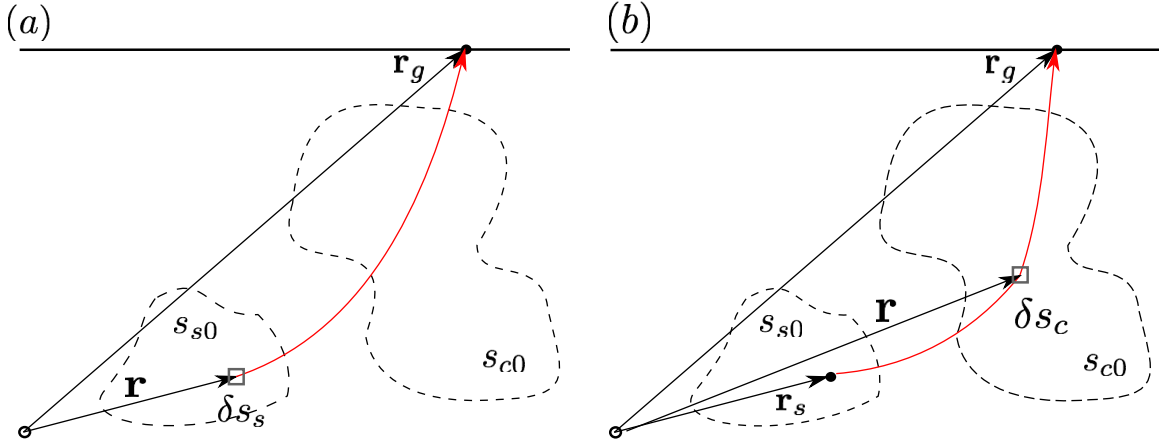
FIG. 3. Schematic illustration of model space with elements δs_s (a) and δs_c (b).

This can then be rearranged to become

$$\frac{\partial P(\mathbf{r}_g, \omega)}{\partial s_c(\mathbf{r})} = \omega^2 G(\mathbf{r}_g, \mathbf{r}, \omega | s_{c0}, s_{s0}) \int d\mathbf{r}_s G(\mathbf{r}, \mathbf{r}_s, \omega | s_{c0}, s_{s0}) s_{s0}(\mathbf{r}_s), \tag{20}$$

which is schematically illustrated in Figure 3b. Knowing the sensitivities, obtaining the gradients is simply a matter of substituting into the following formula; for the derivation refer to Margrave et al. (2011). For the case of the source distribution sensitivity,

$$g_s(\mathbf{r}_g) = \int d\omega \frac{\partial P(\mathbf{r}_g, \omega)}{\partial s_s(\mathbf{r}_g)} \delta P^*(\mathbf{r}_g, \mathbf{r}_s, \omega | s_0), \tag{21}$$

where $g_s(\mathbf{r})$ is the gradient and the * denotes the complex conjugate. The same form is true for the velocity distribution sensitivity.

While it may be easier to visualize the effects of the microseismic inverse problem in frequency domain, the problem may be more tractable in time domain, which is the subject of the next section.

Time domain gradients

In the time domain, the process to obtain the gradient is similar. We start with an acoustic source $s_s(\mathbf{r}, t)$ that has an unknown spatial distribution and an unknown ignition time $t = t_0$. The scalar acoustic medium, which is also unknown, is $s_c(\mathbf{r}) = \frac{1}{c^2(\mathbf{r})}$ as before. In this case, the acoustic wave equation is

$$\left[\nabla^2 - s_c(\mathbf{r}_g) \frac{\partial^2}{\partial t^2} \right] p(\mathbf{r}_g, t | s_c, s_s) = s_s(\mathbf{r}_g, t). \tag{22}$$

As in the frequency domain case, consider the actual source/medium arrangement to be a perturbation on a background arrangement $s_{c0}(\mathbf{r})$, $s_{s0}(\mathbf{r}, t)$. That is,

$$\left[\nabla^2 - s_{c0}(\mathbf{r}_g) \frac{\partial^2}{\partial t^2} \right] p(\mathbf{r}_g, t | s_{c0}, s_{s0}) = s_{s0}(\mathbf{r}_g, t). \tag{23}$$

Similarly, with a delta function source, we have our familiar Green's function represented again:

$$\left[\nabla^2 - s_{c0}(\mathbf{r}_g)\frac{\partial^2}{\partial t^2}\right]g(\mathbf{r}_g, t, \mathbf{r}_s, t_0|s_{c0}, s_{s0}) = \delta(\mathbf{r}_g - \mathbf{r}_s)\delta(t - t_0). \qquad (24)$$

Recall that if we time shift the source time t_0 and the measurement time t both, it will have no effect on the behaviour of the equation. To clean up some of the upcoming stamp collecting, let us subtract t_0 from both t_0 and t in the above equation. We will change our notation as follows:

$$g(\mathbf{r}_g, t, \mathbf{r}_s, t_0|s_{c0}, s_{s0}) = g(\mathbf{r}_g, t - t_0, \mathbf{r}_s, 0|s_{c0}, s_{s0}) \qquad (25)$$
$$= g(\mathbf{r}_g, \mathbf{r}_s, t - t_0|s_{c0}, s_{s0}). \qquad (26)$$

As in the frequency domain case, the field p at $\mathbf{r}_g$ due to a background source s_{s0} in a background medium s_{c0} is a superposition over $\mathbf{r}_s$. However, now that we have time as an added variable, we also have superposition over t_0. Therefore,

$$p(\mathbf{r}_g, t|s_{c0}, s_{s0}) = \int d\mathbf{r}_s \int dt_0 g(\mathbf{r}_g, \mathbf{r}_s, t - t_0|s_{c0}, s_{s0})s_{s0}(\mathbf{r}_s, t_0) \qquad (27)$$

$$= \int d\mathbf{r}_s g(\mathbf{r}_g, \mathbf{r}_s, t|s_{c0}, s_{s0}) * s_{s0}(\mathbf{r}_s, t) \qquad (28)$$

where $*$ denotes convolution. Going back to the wave equation, the perturbed equation with the s_{c0} operator is

$$\left[\nabla_g^2 - s_{c0}(\mathbf{r}_g)\frac{\partial^2}{\partial t^2}\right]g(\mathbf{r}_g, \mathbf{r}_s, t - t_s|s_c) = \delta(\mathbf{r}_g - \mathbf{r}_s)\delta(t - t_s) + \delta s_c(\mathbf{r}_g)\frac{\partial^2}{\partial t^2}g(\mathbf{r}_g, \mathbf{r}_s, t - t_s|s_c)$$
$$\qquad (29)$$

where everything right of the equals sign is the new source term. We can plug this new source term into Eq. 27. The wavefield with this source is really a Green's function so

$$g(\mathbf{r}_g, \mathbf{r}_s, t - t_s|s_c) \approx \int d\mathbf{r}' \int dt' g(\mathbf{r}_g, \mathbf{r}', t - t'|s_{c0})\Big[\delta(\mathbf{r}' - \mathbf{r}_s)\delta(t' - t_s) \qquad (30)$$

$$+ \delta s_c(\mathbf{r}')\frac{\partial^2}{\partial t'^2}g(\mathbf{r}', \mathbf{r}_s, t' - t_s|s_{c0})\Big]. \qquad (31)$$

In order to simplify this expression, we need to side-track and review a special form of convolution. Given

$$f(t - t_0) = \int dt' g(t - t')h(t' - t_0), \qquad (32)$$

let $t'' = t' - t_0$. This means that $t' = t'' + t_0$ and $dt'' = dt'$ so we can transform Eq. 32 into

$$f(t - t_0) = \int dt'' g((t - t_0) - t'')h(t''). \qquad (33)$$

Why is this useful? Looking at to Eq. 30, we need this to sort out the various time variables that combine when we multiply through the Green's function with the second term in the

square brackets. The first term in the expansion is simple because the two delta functions take care of the two integrals. That is,

$$g(\mathbf{r}_g, \mathbf{r}_s, t - t_s | s_c) \approx g(\mathbf{r}_g, \mathbf{r}_s, t - t_s | s_{c0}) \tag{34}$$

$$+ \int d\mathbf{r}' \delta s_c(\mathbf{r}') \int dt' g(\mathbf{r}_g, \mathbf{r}', (t - t_s) - t' | s_{c0}) \left[\frac{\partial^2}{\partial t'^2} g(\mathbf{r}', \mathbf{r}_s, t' | s_{c0}) \right] \tag{35}$$

Now we have all the building blocks to determine the residuals and get the root formula from which we can solve for the two sensitivities by setting individual perturbations to zero. To reiterate, we have the full true wavefield

$$p(\mathbf{r}_g, t | s_c, s_s) = \int d\mathbf{r}_s g(\mathbf{r}_g, \mathbf{r}_s, t | s_c) * s_s(\mathbf{r}_s, t), \tag{36}$$

and the field with the background source/medium:

$$p(\mathbf{r}_g, t | s_{c0}, s_{s0}) = \int d\mathbf{r}_s g(\mathbf{r}_g, \mathbf{r}_s, t | s_{c0}) * s_{s0}(\mathbf{r}_s, t). \tag{37}$$

We define the residual as $\delta p = p(s_c, s_s) - p(s_{c0}, s_{s0})$ so in order to make this work, we need to write Eq. 36 in terms of the background/source medium, as well as the perturbations. That is, in Eq. 36, we substitute $s_c(\mathbf{r}) = s_{c0}(\mathbf{r}) + \delta s_c(\mathbf{r})$ and $s_s(\mathbf{r}) = s_{s0}(\mathbf{r}) + \delta s_s(\mathbf{r})$ to get

$$p(\mathbf{r}_g, t | s_c, s_s) \approx \int d\mathbf{r}_s \int dt_s \Big[g(\mathbf{r}_g, \mathbf{r}_s, t - t_s | s_{c0}) \tag{38}$$

$$+ \int d\mathbf{r}' \delta s_c(\mathbf{r}') \int dt' g(\mathbf{r}_g, \mathbf{r}', (t - t_0) - t' | s_{c0}) \frac{\partial^2}{\partial t'^2} g(\mathbf{r}', \mathbf{r}_s, t' | s_{c0}) \Big] \tag{39}$$

$$\cdot \left[s_{s0}(\mathbf{r}_s, t_s) + \delta s_s(\mathbf{r}_s, t_s) \right]. \tag{40}$$

Expanding this leads to four terms:

$$p(\mathbf{r}_g, t | s_c, s_s) \approx p(\mathbf{r}_g, t | s_{c0}, s_{s0}) + \int d\mathbf{r}_s \int dt_s g(\mathbf{r}_g, \mathbf{r}_s, t - t_s | s_{c0}) \delta s_s(\mathbf{r}_s, t_s)$$

$$+ \int d\mathbf{r}_s \int dt_s \int d\mathbf{r}' \delta s_c(\mathbf{r}') \int dt' g(\mathbf{r}_g, \mathbf{r}', (t - t_s) - t' | s_{c0}) \frac{\partial^2}{\partial t'^2} g(\mathbf{r}', \mathbf{r}_s, t' | s_{c0}) s_{s0}$$

$$+ \int d\mathbf{r}_s \int dt_s \int d\mathbf{r}' \delta s_c(\mathbf{r}') \int dt' g(\mathbf{r}_g, \mathbf{r}', (t - t_s) - t' | s_{c0}) \frac{\partial^2}{\partial t'^2} g(\mathbf{r}', \mathbf{r}_s, t' | s_{c0}) \delta s_s.$$

Knowing that the fourth term will always be zero, we neglect it from here on in.

Now we can write our residuals out and simplify them as follows:

$$\delta p(\mathbf{r}_g, t | s_{c0}, s_{s0}) = p(\mathbf{r}_g, t | s_c, s_s) - p(\mathbf{r}_g, t | s_{c0}, s_{s0}) \tag{41}$$

$$= \int d\mathbf{r}_s g(\mathbf{r}_g, \mathbf{r}_s, t | s_{c0}) * \delta s_s(\mathbf{r}_s, t) \tag{42}$$

$$+ \int d\mathbf{r}' \delta s_c(\mathbf{r}') \int d\mathbf{r}_s O(\mathbf{r}_g, \mathbf{r}', \mathbf{r}_s, t) * s_{s0}(\mathbf{r}_s, t), \tag{43}$$

where

$$O(\mathbf{r}_g, \mathbf{r}', \mathbf{r}_s, t) = g(\mathbf{r}_g, \mathbf{r}', t | s_{c0}) * \frac{\partial^2}{\partial t^2} g(\mathbf{r}', \mathbf{r}_s, t | s_{c0}). \tag{44}$$

Finally we can solve for the sensitivities. Let us solve for the source distribution sensitivity by setting $\delta s_c = 0$ and $\delta s_s(\mathbf{r}_s, t) = \delta s_s(\mathbf{r}^*, t^*)\delta(\mathbf{r}^* - \mathbf{r}_s)\delta(t - t^*)$. Therefore,

$$\delta p_s(\mathbf{r}_g, t | s_{c0}, s_{s0}) \approx g(\mathbf{r}_g, \mathbf{r}^*, t - t^*)\delta s_s(\mathbf{r}^*, t^*), \tag{45}$$

which means

$$\frac{\partial p(\mathbf{r}_g, t)}{\partial s_s(\mathbf{r}^*, t^*)} = g(\mathbf{r}_g, \mathbf{r}^*, t - t^*). \tag{46}$$

Next we repeat this analysis in the case of the velocity model sensitivities where $\delta s_s = 0$ and $\delta s_c(\mathbf{r}') = \delta s_c(\mathbf{r}^*)\delta(\mathbf{r}^* - \mathbf{r}')$:

$$\delta p_c(\mathbf{r}_g, t | s_{c0}, s_{s0}) \approx \delta s_c(\mathbf{r}^*) \int d\mathbf{r}_s O(\mathbf{r}_g, \mathbf{r}^*, \mathbf{r}_s, t) * s_{s0}(\mathbf{r}_s, t). \tag{47}$$

Therefore, the sensitivity is

$$\frac{\partial p(\mathbf{r}_g, t)}{\partial s_c(\mathbf{r}^*)} = \int d\mathbf{r}_s \left[g(\mathbf{r}_g, \mathbf{r}^*, t | s_{c0}) * \frac{\partial^2}{\partial t^2} g(\mathbf{r}^*, \mathbf{r}_s, t | s_{c0}) \right] * s_{s0}(\mathbf{r}_s, t) \tag{48}$$

which in long-form is

$$\frac{\partial p(\mathbf{r}_g, t)}{\partial s_c(\mathbf{r}^*)} = \int d\mathbf{r}_s \int dt' \int dt_s g(\mathbf{r}_g, \mathbf{r}^*, (t - t_s) - t' | s_{c0}) \frac{\partial^2}{\partial t'^2} g(\mathbf{r}^*, \mathbf{r}_s, t' | s_{c0}) s_{s0}(\mathbf{r}_s, t_s). \tag{49}$$

The final step is to write the expression for the gradient based on the sensitivity. For the source-distribution uncertainty, the gradient is given by

$$g_s = -\sum_{r_g, r_s} \int dt \delta P(\mathbf{r}_g, \mathbf{r}_s, t | s_c, s_s) g(\mathbf{r}_g, \mathbf{r}, t - t^* | s_c, s_s). \tag{50}$$

where we see that the gradient is simply given by the sum of the residuals multiplied by the green's function. The companion report (Igonin and Innanen, 2017) shows some numeric examples of what this update ends up looking like.

The velocity model time-domain gradient is given by

$$g_c = -\sum_{r_g, r_s} \int dt \delta P^*(\mathbf{r}_g, \mathbf{r}_s, t | s_c, s_s) \int dt' \int dt_s g(\mathbf{r}_g, \mathbf{r}, (t - t_s) - t' | s_{c0}, s_{s0}) \frac{\partial^2 g(\mathbf{r}, \mathbf{r}_s, t' | s_{c0}, s_{s0})}{\partial t'^2} s_{s0}(\mathbf{r}_s, t_s). \tag{51}$$

This concludes the time domain derivation of the sensitivities of the source distribution and velocity model.

Frequency-domain Hessian

Since we have a lot of interest in the cross-talk between the source distribution and velocity model uncertainties, the Hessian is a key component of the analysis. It is the diagonal terms of the Hessian that contain the shared terms and ultimately, will be the gate

to address cross-talk that may exist. The Hessian is a 2×2 block matrix, as a virtue of having two sensitivities. It can be written as

$$H(\mathbf{r}, \mathbf{r}') = \begin{bmatrix} H_1(\mathbf{r}, \mathbf{r}') & H_2(\mathbf{r}, \mathbf{r}') \\ H_3(\mathbf{r}, \mathbf{r}') & H_4(\mathbf{r}, \mathbf{r}') \end{bmatrix}. \tag{52}$$

We will now proceed to derive the four components of the Hessian. Since the Hessian is by definition symmetric, we need only find the expressions of three terms. Additionally, we will refrain to the Gauss-Newton Hessian, which is given by

$$H(\mathbf{r}, \mathbf{r}') = \int d\omega \frac{\partial P}{\partial s_0(\mathbf{r}')} \frac{\partial P^*}{\partial s_0(\mathbf{r})}, \tag{53}$$

where the $s_0(\mathbf{r})$ refers to either the $s_c(\mathbf{r})$ or $s_s(\mathbf{r})$, depending on the case.

The first term, $H_1(\mathbf{r}, \mathbf{r}')$ is due to only the source distribution sensitivities and due to the simple nature of the sensitivity is given by

$$H_1(\mathbf{r}, \mathbf{r}') = \int d\omega G(\mathbf{r}_g, \mathbf{r}', \omega | s_{c0}, s_{s0}) G^*(\mathbf{r}_g, \mathbf{r}, \omega | s_{c0}, s_{s0}). \tag{54}$$

The second and third terms are given by

$$H_2(\mathbf{r}, \mathbf{r}') = \int d\omega G(\mathbf{r}_g, \mathbf{r}', \omega | s_{c0}, s_{s0}) \omega^2 G^*(\mathbf{r}_g, \mathbf{r}, \omega | s_{c0}, s_{s0}) \int d\mathbf{r}_s G(\mathbf{r}, \mathbf{r}_s, \omega | s_{c0}, s_{s0}) s_{s0}(\mathbf{r}_s). \tag{55}$$

Finally, the fourth and longest term, which is due to the velocity model sensitivities alone, is given by

$$H_4(\mathbf{r}, \mathbf{r}') = \int d\omega \, \omega^4 G(\mathbf{r}_g, \mathbf{r}', \omega | s_{c0}, s_{s0}) G^*(\mathbf{r}_g, \mathbf{r}, \omega | s_{c0}, s_{s0}) \cdot \tag{56}$$

$$\int d\mathbf{r}_s s_{s0}(\mathbf{r}_s)^2 G(\mathbf{r}', \mathbf{r}_s, \omega | s_{c0}, s_{s0}) G^*(\mathbf{r}, \mathbf{r}_s, \omega | s_{c0}, s_{s0}). \tag{57}$$

While these formulae are informative, one of the best ways to explore their meaning is with an analytic example, which is the subject of the following section.

ANALYTIC EXAMPLE

Let us begin the analytic example by describing our model space. Consider a 1D universe in the variable z. At the surface $z = z_g = 0$, we have a geophone, and at some depth z_s we have a source. The wavefront generated at z_s travels through a homogeneous medium $s(z) = s_0$, as shown in Figure 4. Since we would like to explore the behaviour of our new source distribution sensitivity, we will maintain that the velocity is known accurately. In order to study the Hessian as well, we will focus on the frequency domain. The true data is a one-way Green's function given by

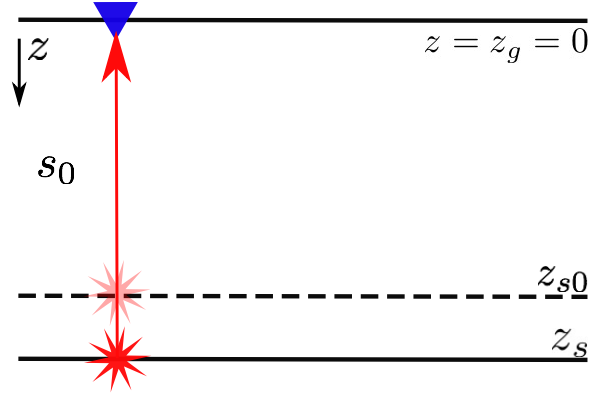$$D(z_s, z_g, \omega) = \frac{e^{ik_0 z_s}}{i2k_0}, \tag{58}$$

FIG. 4. Schematic illustration of model space for the analytic example.

where $k_0 = \frac{w}{c_0}$. Let's say that the true depth is unknown and we guess the starting point to be z_{s0}. Therefore, the modelled data is given by

$$P(z_{s0}, z_g, \omega) = \frac{e^{ik_0 z_{s0}}}{i2k_0}. \tag{59}$$

The next step is to calculate the complex conjugate of the residuals, which is given by

$$\delta P^*(z_g, z_s, \omega) = \frac{-e^{-ik_0 z_{s0}} + e^{-ik_0 z_s}}{i2k_0}. \tag{60}$$

The source distribution sensitivity is simply the Green's function, so

$$\frac{\partial P(z_g, \omega)}{\partial s_s(z)} = G(z_g, z, \omega | s_{c0}, s_{s0}) = \frac{e^{ik_0 z}}{i2k_0}. \tag{61}$$

This can be substituted into the expression for the gradient:

$$g_s(z) = \int d\omega \frac{\partial P(z_g, \omega)}{\partial s_s(z)} \delta P^*(z_g, z_s, \omega | s_0) \tag{62}$$

$$= \int d\omega \left[\frac{-e^{ik_0(z-z_{s0})}}{(i2k_0)^2} + \frac{e^{ik_0(z-z_s)}}{(i2k_0)^2} \right] \tag{63}$$

Based on the form of this equation, and recalling what delta and step functions look like, this is going to be a ramp function. There are a few things we can do to solve this integral:

→ Discretize the integral, keeping only the low frequencies and effectively remove the integral.

→ Determine the Fourier-transform (FT) expression for a ramp function.

→ Expand the exponential terms as a Taylors' series and keep only the low orders for integration.

Let us try the second approach. Recall that the FT for the delta function is given by

$$\delta(z - z_0) = \frac{1}{2\pi} \int dk_z e^{ik_z(z-z_0)}.$$ (64)

Likewise, we have the Heaviside, or step, function given by

$$\delta^{(-1)}(z - z_0) = \frac{1}{2\pi} \int dk_z \frac{e^{ik_z(z-z_0)}}{ik_z} \equiv \theta(z - z_0).$$ (65)

Therefore, in a similar way, we can define

$$\delta^{(-2)}(z - z_0) = \frac{1}{2\pi} \int dk_z \frac{e^{ik_z(z-z_0)}}{(ik_z)^2} \equiv \psi(z - z_0),$$ (66)

where $\psi(z - z_0)$ is the ramp function.

Applying this to Eq. 63 and performing a change in variables to make the integral in terms of dk_0, we obtain the gradient

$$g_s(z) = \frac{c_0\pi}{2} \left[-\psi(z - z_{s0}) + \psi(z - z_s) \right].$$ (67)

Figure 5a shows what these two ramp functions look like in combination for the case of $z_{s0} < z_s$. If we had instead said that the background velocity was also incorrect, with k_1 being the new wavenumber in this case, the gradient would have been

$$g_s(z) = \frac{c_0\pi}{2} \left[-\psi(z - z_{s0}) + \psi(z - \frac{k_1}{k_0}z_s) \right].$$ (68)

Figure 5b shows the effect on the gradient depending on whether $k_1 < k_0$ or vice versa. What this shows is that if the background medium is not known accurately, the source-term gradient will attempt to move the source deeper or shallower to account for it.
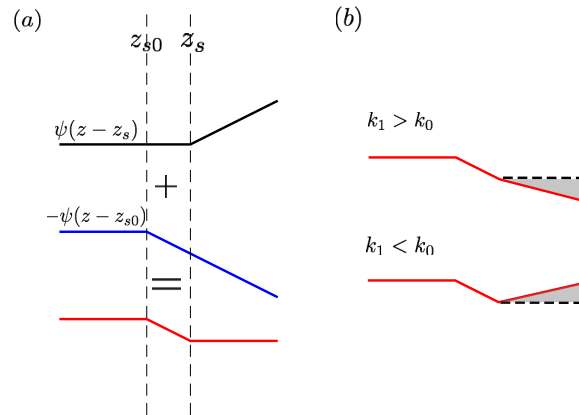

FIG. 5. Illustration of combined ramp functions for Eq. 68 and 69. (a) In the case of a known background medium; (b) in the case of a faster or slower true background medium.

Next, we can determine the gradient for the velocity model sensitivity. The sensitivity is given by

$$\frac{\partial P(z_g, \omega)}{\partial s_c(z)} = \omega^2 \frac{e^{ik_0z}}{i2k_0} \int dz_s \frac{e^{ik_0|z-z_s|}}{(i2k_0)^2} s_{s0}(z_s).$$ (69)

If we have a point source at $z_s = z_{s0}$, the integral goes away and we have

$$\frac{\partial P(z_g, \omega)}{\partial s_c(z)} = -\frac{1}{(i2k_0)^2}\omega^2 e^{ik_0(2z - z_{s0})}, \tag{70}$$

After substituting this into the formula for the gradient, the resulting expression can be simplified by transforming the integral to be with respect to $d2k$ and using Eq. 65, the definition of the Heaviside function. Therefore, the gradient becomes

$$g_c(z) = \frac{c_o^3 \pi}{4} \left[-\theta(z - z_{s0}) + \theta(z - (\frac{z_{s0}}{2} + \frac{z_s}{2})) \right]. \tag{71}$$

Examining the form of this closely, we see that we have two step functions of the same amplitude, but in two different locations. If we have $z_s > z_{s0}$, the two step functions in combination will create a box function that starts at z_{s0} and ends at $\frac{z_{s0}}{2} + \frac{z_s}{2}$. Another observation is that even though the velocity model is known, the gradient creates a non-zero update in the region where the source is uncertain.

Finally, using Eq. 53 through 58, the Hessian can be constructed to further analyze the cross-talk and begin to see what the Gauss-Newton update would look like for the first iteration. After some math, the Hessian can be summarized as

$$H(\mathbf{r}, \mathbf{r}') = \begin{bmatrix} H_1(z, z') & H_2(z, z') \\ H_3(z, z') & H_4(z, z') \end{bmatrix} = \begin{bmatrix} \frac{c_0\pi}{2}\psi(z' - z) & \frac{c_0\pi}{4}\theta(z' - 2z + z_{s0}) \\ \frac{c_0\pi}{4}\theta(z' - 2z + z_{s0}) & \frac{c_0\pi}{16}\delta(z' - z) \end{bmatrix}. \tag{72}$$

Recalling that the Gauss-Newton update is given by

$$\delta s = -H^{-1}g, \tag{73}$$

we can imagine determining the inverse of each element in Eq. 73. Then consider multiplying this inverse Hessian first by $g_s(z)$. Interestingly enough, we end up with terms like $\psi^{-1}(...)\psi(...)$ and $\theta^{-1}(...)\theta(...)$. Therefore, the step may end up resembling a delta function. Next, consider the velocity model update. In that case, we have terms like $\theta^{-1}(...)\psi(...)$ and $\delta(...)\theta(...)$, which may lead to the update resembling a step. In summary, the Hessian may in fact give the update the correct character we would expect; delta functions for source position, and steps for velocity model layers.

DISCUSSION AND FUTURE WORK

Since this project is still in its infancy, the scope of the future work is quite broad. The first order of business is to implement the new gradient in the framework of an existing FWI code. This has been attempted in the companion report, Igonin and Innanen (2017).

The Hessian is the key to studying the cross-talk between these two parameters and a major focus in the future is to characterize and propose solutions to this cross-talk.

The ultimate goal is to be able to apply an iterative FWI scheme for these two parameters with real data. Due to the complex wave nature of microseismic data, the first attempt would be made on physical modelling lab data, and only later would be applied to a real microseismic dataset owned by the Microseismic Industry Consortium (MIC). An extension to 3D elastic media would also be necessary in this case, which would yield more complex expressions for the gradient and Hessian.

CONCLUSIONS

In this report, sensitivities for a source and velocity model uncertainty were derived as a first step toward a transmission-based iterative FWI scheme to solve for both source location and velocity model. The sensitivities were derived in both time and frequency domain. Additionally, the Hessian was derived in the frequency domain. To explore the behaviour of the gradient and Hessian, an analytic example in frequency domain was undertaken. While this led to a few speculative insights, this research project has a broad scope of future work with the goal of being eventually applied to real data, both laboratory and field in origin.

ACKNOWLEDGEMENTS

REFERENCES

Alkhalifa, T., and Plessix, R., 2014, A recipe for practical full-waveform inversion in anisotropic media: An analytic parameter resolution study: Geophysics, **79**, R91–R101.

Artman, B., Podladtchikov, I., and Witten, B., 2010, Source location using time-reverse imaging: Geophysical Prospecting, **58**, No. 5, 861–873.

Bardainne, T., and Gaucher, E., 2010, Constrained tomography of realistic velocity models in microseismic monitoring using calibration shots: Geophysical Prospecting, **58**, No. 5, 739–753.

Butzer, S., Kurzmann, A., and Bohlen, T., 2013, 3D elastic full-waveform inversion of small-scale heterogeneities in transmission geometry: Geophysical Prospecting, **61**, No. 6, 1238–1251.

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

Igonin, N., and Innanen, K., 2017, Applications of FWI to the microseismic source problem: CREWES Research Report, **29**.

Jarillo, M., and Tsvankin, I., 2017, Waveform inversion for microseismic velocity analysis and event location in vti media: Geophysics, **82**, No. 4, WA95–WA103.

Kamath, N., and Tsvankin, I., 2014, Elastic full-waveform inversion of transmission data in 2D VTI media: SEG Technical Program Expanded Abstracts 2014, 1157–1161.

Long, G., Li, X., Zhang, M., and Zhu, T., 2009, Visco-acoustic transmission waveform inversion of velocity structure in space-frequency domain: Earthquake Science, **22**, 45–52.

Margrave, G., Yedlin, M., and Innanen, K., 2011, Full waveform inversion and the inverse hessian: CREWES Research Report, **23**, 77.1–77.13.

Maxwell, S., Bennett, L., and Walsh, J., 2010, Anisotropic velocity modeling for microseismic processing: Part 1–impact of velocity model uncertainty: SEG Technical Program Expanded Abstracts 2010, 2130–2134.

Pan, W., Innanen, K., Margrave, G., Fehler, M., Fang, X., and Li, J., 2016, Estimation of elastic constants for HTI media using Gauss-Newton and full-Newton multiparameter full-waveform inversion: Geophysics, **81**, R275–R291.

Sethi, H., and Shekar, B., 2017, Full-waveform inversion for microseismic events using sparcity constraints: SEG Technical Program Expanded Abstracts 2017, 2888–2892.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, No. 8, 1259–1266.

Virieux, J., and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, No. 6, WCC127–WCC152.

Waldhauser, F., and Ellsworth, W. L., 2000, A double-difference earthquake location algorithm: Method and application to the northern hayward fault, california: Bulletin of the Seismological Society of America, **90**, No. 6, 1353.